University of Alberta

Conformational Studies of Small Peptides by ^{1}H Homonuclear 2D NMR and Molecular Modeling.

by

Mark Jason Miskolzie

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science.

Department of Chemistry

Edmonton, Alberta

Spring, 2003

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-82314-8

Canada

University of Alberta

Library Release Form

Name of Author: Mark Jason Miskolzie

Title of Thesis: Conformational Studies of Small Peptides by ^{1}H Homonuclear 2D NMR and Molecular Modeling.

Degree: Master of Science.

Year this Degree Granted: 2003

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Mark Jason Miskolzie
8230 94 ave
Fort Saskatchewan, Alberta
Canada
T8L 2V2

Date: Jan 10, 2003.

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Conformational Studies of Small Peptides by ^{1}H Homonuclear 2D NMR and Molecular Modeling." submitted by Mark Jason Miskolzie in partial fulfillment of the requirements for the degree of Master of Science.

Dr. George Kotovych

Dr. Roderick E. Wasylishen

Dr. Leo Spyracopoulos

Date: Dec 16, 2002

Abstract

The conformations of three peptides were determined by ^{1}H two-dimensional NMR and molecular modeling. The peptides that were studied were B-10324, a bradykinin B_1 receptor antagonist; neuropeptide F, a peptide isolated from the parasitic tapeworm *Moniezia expansa*; and neuropeptide AF, a pain modulating and anti-opiate neuropeptide isolated from humans. The secondary structure of B-10324 in 60%/40% CD_3OH/H_2O was determined to have a distorted type II β-turn between residues Pro^{2}-Cpg^{5} with a stabilizing hydrogen bond between the same residues, where Cpg is cyclopentylglycine. Neuropeptide F displayed an amphipathic α-helix in 60%/40% CD_3OH/H_2O from residues Lys^{14}-Ile^{31}, and the results were compared to porcine neuropeptide Y bound to dodecylphosphocholine micelles. Finally, the secondary structure of neuropeptide AF was determined in two different solvents, namely 50%/50% TFE/H_2O and sodium dodecylsulphate micelles. Under both solvent conditions the secondary structure of NPAF consisted of an α-helix approximately between, Ser^{7}-Ala^{14}. In the sodium dodecylsulphate micelle solution, the C-terminal tetrapeptide was structured, forming a type I β-turn.

Acknowledgements:

First and foremost I would like to thank my supervisor, Dr. George Kotovych, for the years of guidance, patience, support, and especially for the opportunity to study as a graduate student.

I would also like to thank Dr. Tom Nakishima, for many meaningful conversations and hours of help with the spectrometer throughout my degree.

I would like to thank Dr. John M. Stewart, for the donation of interesting medicinally important peptides, which were invaluable for this thesis.

Finally, I would like to thank my loving parents, Will and Joanne, for years of support. My dearest Penny for her support, patience, and friendship which was essential throughout my time as a graduate student.

The Canadian National High Field NMR Centre (NANUC) for their assistance and use of the facilities. Operation of NANUC is funded by the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada and the University of Alberta.

The author would like to thank the University of Alberta, and the Natural Sciences and Engineering Research Council of Canada for financial assistance.

Table of Contents

Chapter	Page
Chapter 1. Introduction	1
Introduction	2
Protein Structure	2
Structure Determination	5
The NMR Phenomenon	5
2D NMR and Peptide Structure Determination	10
Chemical Shift	13
Experimental Distances	14
Molecular Modeling	15
References	26
Chapter 2. Correlation of Secondary Structures of Bradykinin B_1 Receptor Antagonists with Their Activity.	28
Introduction	29
Materials and Methods	33
NMR Sample Preparation	33
NMR Spectroscopy	33
Molecular Modeling	35
Results and Discussion	36
References	45
Chapter 3. The NMR-derived Conformation of Neuropeptide F from *Moniezia expansa*.	47

Introduction 48

Methods and Materials 51

NMP. Sample Preparation 51

NMR Spectroscopy 51

Molecular Modeling 53

Results and Discussion 54

References 67

Chapter 4. The NMR-Derived Conformation of Neuropeptide AF, an Orphan G-Protein Coupled Receptor Peptide. 70

Introduction 71

Materials and Methods 72

Materials 72

Methods 73

The NMR Sample 73

TFE/H_2O Solution 73

SDS Solution 73

Spin Label Studies 74

NMR 74

Molecular Modeling 76

Results and Discussion 77

Spectroscopy 77

Npaf in TFE 78

SDS 81

Spin Label Studies 86

Summary and Conclusions 87

References 105

Chapter 5. Concluding Chapter 109

Conclusions 110

Future Work 112

B-10324 112

NPF 112

NPAF 113

References 115

List of Tables

Table		**Page**
Table 1-1:	The idealized dihedral angles for various turns found in protein structure.	19
Table 2-1:	The proton chemical shifts, coupling constants and NH temperature coefficients of B-10324.	41
Table 3-1:	The ^1H chemical shifts of NPF in CD_3OH/H_2O 60%/40% by volume, referenced to CHD_2OH.	59
Table 3-2:	Structural statistics for the calculated 20 lowest energy conformers of NPF.	61
Table 4-1:	The chemical shifts (ppm) and coupling constants (Hz) for NPAF in 50%/50% TFE/H_2O at 20°C.	90
Table 4-2:	The chemical shifts (ppm) and coupling constants (Hz) for NPAF in 243mM SDS at 35°C.	91
Table 4-3:	The structural statistics for NPAF in TFE/H_2O.	92
Table 4-4:	The structural statistics for NPAF in SDS.	93

List of Figures

Figure **Page**

Figure 1-1: The polypeptide backbone showing the defining dihedral angles. 20

Figure 1-2: Idealized type I and type II beta turns. 21

Figure 1-3: The TOCSY pulse sequence with presaturation of the water peak for solvent suppression. 22

Figure 1-4: The pulse sequence for a GCOSY with presaturation of the water peak for solvent suppression. 23

Figure 1-5: The pulse sequence for a NOESY experiment with presaturation of the water peak for solvent suppression. 24

Figure 1-6: The expected NOE's and coupling constants for sections of regular secondary structure. 25

Figure 2-1: The primary structures of the bradykinin B_1 receptor antagonists B-9858, B-10148, and B-10324. 42

Figure 2-2: The αH-NH region of the 200ms NOESY at 800MHz and 5°C observed for B-10324. 43

Figure 2-3: The stereo view of the 10 lowest energy conformers of B-10324 showing the 2-5 type II β-turn. 44

Figure 3-1: Comparison of NPF with porcine NPY and PYY sequences. 62

Figure 3-2: Part of the fingerprint region of the 800MHz, 150ms mixing time, NOESY spectrum for NPF in 60%/40% CD_3OH/H_2O at 298K. 63

Figure 3-3: The NOE summary and secondary structure of NPF. 64

Figure 3-4: The stereo view of the 20 lowest energy conformers of NPF showing the best fit superposition of the backbone heavy atoms along the helical region, from K^{14} to I^{31}. 65

Figure 3-5: The twenty lowest energy structures of NPF showing the amphipathic alpha helix. 66

Figure 4-1: Part of the NH-Hα region of the 600MHz NOESY with a 150ms mixing time for NPAF in 50%/50% TFE/H_2O. 94

Figure 4-2: The summary of important NOE data for NPAF in 50%/50% TFE/H_2O at 20°C. 95

Figure 4-3: The H_α chemical shift index of NPAF in TFE/H_2O. 96

Figure 4-4: The stereo view of the twenty lowest energy structures of NPAF in TFE/H_2O showing the best fit superposition of residues Asn^6 to Ala^{13}. 97

Figure 4-5: The chemical shift index for NPAF in SDS. 98

Figure 4-6: Part of the NH-Hα region of the 150ms mixing time 800MHz NOESY spectrum of NPAF in 243mM SDS taken at 35°C. 99

Figure 4-7: The summary of important NOE data for NPAF in 243mM SDS at 35°C. 100

Figure 4-8: The stereo view of the twenty lowest energy structures of NPAF in 243mM SDS showing the best fit superposition of residues Ser^{7} to Ala^{14}. 101

Figure 4-9: One-dimensional spectra of NPAF in the SDS solution in the presence and in the absence of the spin label 5-doxyl-stearic acid. 103

Figure 4-10: The percentage of signal intensity remaining after the addition of the 5-DS for HN-Hα TOCSY cross peaks. 104

List of Abbreviations

Amino Acids

The twenty naturally occurring amino acids.

A - Ala - Alanine

C - Cys - Cysteine

D - Asp - Aspartic Acid

E - Glu - Glutamic Acid

F - Phe - Phenylanine

G - Gly - Glycine

H - His - Histidine

I - Ile - Isoleucine

K - Lys - Lysine

L - Leu - Leucine

M - Met - Methionine

N - Asn - Asparagine

P - Pro - Proline

Q - Glu - Glutamine

R - Arg - Arginine

S - Ser - Serine

T - Thr - Threonine

V - Val - Valine

W - Trp - Tryptophan

Y - Tyr - Tyrosine

Abnormal Amino Acids

Cpg -	cyclopentylglycine
F5c -	2,3,4,5,6-pentafluorocinnamoyl
F5F -	2,3,4,5,6-pentafluorophenylalanine
Hyp -	*trans*-4-hydroxy-L-proline
Igl -	α-(2-indanyl)glycine
Oic -	(3aS, 7aS)-octahydroindole-2-carboxylic acid
Thi -	beta-(2-thienyl)-L-alanine
Tic -	1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid

Peptides

BK -	bradykinin
NPF -	neuropeptide F
NPY -	neuropeptide Y
PP -	pancreatic polypeptide
PYY –	peptide YY
NPAF -	neuropeptide AF
NPFF -	neuropeptide FF

Chemicals

DSS -	2,2-dimethyl-2-silapentane-5-sulfonic acid
TMS -	tetramethylsilane
5-DS -	5-doxyl-stearic acid
SDS -	sodium dodecylsulphate-d_{25}
TFE -	2,2,2-trifluoroethanol-d_3

DPC - dodecylphosphocholine

Experimental

NMR - nuclear magnetic resonance

TOCSY - total correlation spectroscopy

HOHAHA- Homonuclear Hartman Hahn

NOESY - nuclear Overhauser effect spectroscopy

COSY - correlation spectroscopy

DQFCOSY - double quantum filter COSY

GCOSY - gradient enhance COSY

HSQC - hetronuclear single quantum coherence

NOE - nuclear Overhauser effect

RMSD - root mean squared deviance

ppm - parts per million

ppb - parts per billion

FT- Fourier transform

Hz- Hertz

Chapter 1.

Introduction

Introduction

Proteins and peptides serve a variety of physiologically important roles, from messenger hormone peptides to large catalytic enzymes and muscle complexes. Their function is often dictated by their structure, and a close examination of structural characteristics can yield a wealth of information about the function of the protein. Often the goal of structural studies and the goal of this thesis is to determine the biologically relevant conformation of the peptide or protein of interest to provide insights into their functional properties.

Protein Structure

Protein structure (1) is divided into a hierarchy beginning with the protein's amino acid sequence or primary structure. Structurally, within the polypeptide sequence is the polypeptide backbone. The backbone is essentially a repeating series of heavy atoms that define the backbone, consisting of the amide nitrogen, the alpha carbon, and the carbonyl carbon ($-N_i-C_{\alpha,i}-C'_i-$) of each amino acid as shown in Fig. 1-1. Each amino acid acts as a building block of the polypeptide chain and is linked via a peptide bond, or amide bond, between the carbonyl carbon of amino acid i and amide nitrogen of the following amino acid, i+1. This bonding pattern has partial double bond character, and forms a defining dihedral angle of the polypeptide backbone, the ω dihedral angle, represented by the $C_{\alpha,i}-C'_i-N_{i+1}-C_{\alpha,i+1}$ atoms. Owing to the partial double bond character of this bond, there is a high-energy barrier to rotation about this bond (~20kcal/mol) (2); thus ω is essentially fixed in a flat planar conformation. The preferred conformation is in fact the *trans* conformation or a dihedral angle of 180° for ω. An alternate

conformation for the ω dihedral angle does exist at 0° for the *cis* conformation, generally found at proline residues. In fact, up to 30% of proline residues can exist in the *cis* conformation.

The conformational space of the protein backbone is limited by the rotatable dihedral angles of the protein. There are really only two rotatable dihedral angles within the backbone per amino acid, namely ϕ and ψ, as ω is essentially fixed at 180°. The defining atoms for the ϕ and ψ dihedral angles are C'_{i-1}-N_i-$C_{\alpha,i}$-C'_i and N_i-$C_{\alpha,i}$-C'_i-N_{i+1}, respectively. With essentially two rotatable dihedral angles per amino acid, the allowed conformational space of the protein backbone is limited. The preferred conformational space has been defined, and can be shown graphically in a Ramachandran plot (3,4). The result is that there are a limited number of local three dimensional structures, or secondary structures, that a polypeptide chain can adopt. Secondary structure elements are many and include, but are not limited to, helices, beta sheets, beta turns and random coil sections. These elements of secondary structure elements can be defined wholly by their dihedral angles.

There are two kinds of helices that are found in peptide and protein secondary structure, known as alpha and 3_{10} helices. An alpha helix contains 3.6 residues per turn and, by definition, the ϕ, ψ dihedral angles are -57° and -47°, respectively. On the other hand, a 3_{10} helix contains 3 residues per turn and the defining ϕ,ψ dihedral angles are -60° and -30°, respectively. Short range contacts help to stabilize helices; the most notable of these are hydrogen bonds. Regular patterns are often observed between residues in the i,i+4 positions for alpha

helices and in the i,i+3 positions for 3_{10} helices. The hydrogen bonding donor and acceptor atoms for alpha helices are, for example, the carbonyl oxygen of the i^{th} residue, which serves as the hydrogen bond acceptor, and the amide hydrogen of the i+4 residue, which serves as the hydrogen bond donor ($C{=}O_i$ - HN_{i+4}). Alpha helices are one of the most common secondary structural elements within proteins and peptides.

Beta sheets are a secondary structural feature that are more common in large proteins than small peptides. There are two types of beta sheets that are customarily observed: the parallel and anti-parallel beta sheets. The parallel beta sheet has a parallel arrangement of the polypeptide chain; there is a parallel arrangement of the N- and C-termini. Conversely, the anti-parallel beta sheet has an anti-parallel arrangement of the polypeptide strands; the N- and C-termini have opposite polarity. Long range hydrogen bonds between amino acid pairs across the polypeptide chains act to stabilize beta sheets. Characteristic ϕ,ψ dihedral angles for a parallel beta sheet are -119° and 113°, respectively. The characteristic ϕ,ψ dihedral angles for an anti-parallel beta sheet are -139° and 135°, respectively.

There are many types of turns, which are common to small peptides. Beta turns, or reverse turns, are generally four amino acid turns that reverse or change the direction of the polypeptide chain, and are defined by the dihedral angles of the central two residues. A stabilizing hydrogen bond is often present, but not a prerequisite, for a beta turn. The hydrogen bonding acceptor and donor atoms are, typically, the carbonyl oxygen of residue i or the first amino acid in the turn, and

the amide hydrogen of the i+3 or last amino acid of the turn, respectively. Table 1-1 lists some of the various types of beta turns and defining dihedral angles that are encountered in protein and peptide structures. The inverse of the beta turns is also given. By far the most common beta turns are type I and type II (1,5) shown in Fig. 1-2. Gamma turns, also given in Table 1-1, are three residue turns that are defined by the dihedral angles of the central amino acid.

Finally, random coil sections of the polypeptide chain are usually ill defined and are just that, random.

The protein tertiary structure describes how the elements of secondary structure come together to form the overall three-dimensional fold of the protein. Beyond tertiary structure is quaternary structure which describes how protein subunits come together to create a larger complex.

Structure Determination

There are many methods that can be used to probe the structure of a peptide in solution, but few are more suited to this task than NMR. To meet this end, NMR can probe the structure and dynamics of a protein in its native aqueous environment and yield structural information at the atomic level. The information gained from NMR spectroscopy can be used to determine the tertiary structure of the protein on a level only rivaled by x-ray crystallography (6,7).

The NMR Phenomenon

The phenomenon of NMR is based on the fact that nuclei have a quantum mechanical property known as spin angular momentum; the spin quantum number is denoted by the letter I. The presence of nuclear spin and charge indicates that

the nucleus has a magnetic moment. The magnetic moment is a vector quantity that has both magnitude and direction; the magnitude is given by,

[1] $\mu_m = \gamma\hbar(I(I+1))^{1/2}$

where γ is the magnetogyric ratio, $\hbar$ is Plank's constant divided by 2π, and I is the spin quantum number. Of course, from the above relation it is clear that nuclei with a spin of 0 will be inactive with respect to the basic NMR phenomenon. When a nucleus is placed in a magnetic field, the z-component or direction of the magnetic moment is given by,

[2] $\mu_z = \gamma\hbar m_I$

where m_I is the magnetic quantum number (m_I = -I,(-I+1), (-I+2), …I) and indicates that there are a number of quantized states or orientations for the magnetic moment. Also, the magnetic moment has an associated energy when placed in a homogenous static magnetic field, B_o, and this value is given by the z-component of the magnetic moment,

[3] $E = -\mu_z B_o = -\gamma\hbar m_I B_o$

Equation 3 indicates that there are discrete energy levels for a nucleus in a magnetic field. For a nucleus with a spin of 1/2, there are only two energy levels corresponding to the magnetic quantum numbers m_I=1/2 and m_I= -1/2. These states are also known as the α and β states, respectively, with energies given below.

[4] $E_{ml=1/2} = -1/2\gamma\hbar B_o$ $\qquad$ $E_{ml=-1/2} = 1/2\gamma\hbar B_o$

Nuclei with positive magnetogyric ratios, such as ^1H and ^{13}C, Equation 4 indicates that the α state has lower energy. In the basic NMR experiment we are interested

in transitions between these energy levels, and invoking the Bohr frequency condition, the characteristic frequency, or resonance frequency, of transition for a nucleus in a magnetic field is given by

[5] $\nu = \gamma B_o/2\pi$

where ν is the frequency of transition given in Hz. This is also known as the resonance condition.

Equation 5 indicates that for a spin 1/2 nucleus the NMR spectrum should consist of a single resonance signal for all molecules. However, this is not the case as when a nucleus is placed in a magnetic field, the electrons of the nucleus circulate and induce a small magnetic field that opposes the direction of the static magnetic field. The induced magnetic field is dependent on the environment of the nucleus and the basic resonance condition becomes,

[6] $\nu = \gamma B_o(1-\sigma)/2\pi$

where σ is the shielding constant. The direct result of shielding is that a NMR spectrum consists of several resonance lines characteristic of the environments of the individual nuclei of a molecule, or the chemical shift (δ) which is given by Equation 7 and is measured in parts per million (ppm).

[7] $\delta = ((\nu_{sample} - \nu_{ref})/\nu_{ref})10^6$

The subscripts in Equation 7 denote the resonance frequencies of the sample and a reference compound such as TMS. This is the foundation of the NMR phenomenon.

The second feature of interest in the NMR phenomenon is scalar or spin-spin coupling. The electrons of a nucleus transfer information about their

environment to the other spins in the molecule through the covalent bonding network within the molecule. Scalar couplings, measured in Hertz, manifest themselves in an NMR spectrum as a splitting of resonance signals and are independent of magnetic field strength. Coupling constants can be related to molecular geometry via the Karplus equation (8), shown below.

[8] $^3J = A\cos^2\theta + B\cos\theta + C$

where A, B, and C are empirically derived constants relevant to the coupling constant of interest, J is the coupling constant measured in Hz, the index of 3 refers to the number of bonds over which the coupling constant is measured, and θ is the value of the intervening angle, a dihedral angle. For proteins, the $^3J_{H\alpha HN}$ coupling constant is particularly important since this quantity can be related directly to the ϕ dihedral angle, a defining dihedral angle of the backbone. Consequently, these data can yield direct structural information about the polypeptide backbone. The Karplus equation has been modified to the form,

[9] $^3J_{H\alpha HN} = A\cos^2(\phi - 60^o) + B\cos(\phi - 60^o) + C$

where the previous definitions apply. The values of A, B, and C can be obtained from several sources, reference 9 lists some sources.

For large macromolecules such as proteins and peptides there are simply too many spins for a simple 1D spectrum to be instructive; spectral overlap of resonance signals impedes assignment and structural determination, so NMR spectra of higher dimensionality are required to remove the redundancy. In this thesis small peptides at natural isotope abundance were studied by homonuclear ^1H 2D techniques and, as a result, the emphasis will be placed on homonuclear 2D

NMR experiments. 2D NMR was first proposed by Jeener in 1971 (10) and was preceded by the advent of pulsed FT NMR.

For a clear picture of pulsed FT NMR, one must consider the vector formalism for NMR. Consider an ensemble of spins in the static magnetic field (B_o) in the laboratory frame of reference, which is simply the Cartesian coordinate system with B_o aligned with the positive z-direction. The net magnetization vector at equilibrium will be aligned along the z-direction collinear with the magnetic field. Application of an oscillating radio frequency pulse along the positive x direction perturbs the spins so that the net magnetization is tipped away from the z-direction by some angle θ. If the radio frequency pulse is applied long enough the net magnetization will be placed entirely in the xy-plane and will precess in the xy plane at a frequency referred to as the Larmor frequency. Equation 5 also gives the value of the Larmor frequency (in terms of a radial frequency, the Larmor freqency is $\omega_o=2\pi\nu$). If we consider the frame of reference as rotating at the Larmor frequency and in the same direction as the spins (the rotating frame of reference), then the net magnetization vector will appear static along the negative y direction. In addition, since the net magnetization vector has been rotated by an angle of 90^o away from the z-axis, the rf pulse is referred to as a 90^o pulse. Referring to the laboratory frame of reference once again, the rotating magnetic vector in the static magnetic field will induce an electrical current in the receiver coil. Relaxation processes will restore the net magnetization back to the equilibrium state and the induced voltage will be attenuated in the receiver coil. This signal, known as the free induction decay, is collected as a function of time

and contains all of the relevant NMR information. Fourier transformation converts the time domain signal to the frequency domain, yielding the NMR spectrum.

Two-dimensional NMR utilizes a series of pulses and delays to achieve the desired result, which is best interpreted using the product operator formalism (11). The 2D pulse sequence is divided into 4 periods known as preparation, evolution, mixing and detection. The preparation period consists of a delay to establish thermal equilibrium of the spins followed by one or more pulses to establish the desired coherence. The evolution time (t_1) is a delay that allows the coherences to evolve to a desired state, the mixing period (τ_m) allows for the coherences to transfer between spins, and can be made up of a series of pulses or a delay. Finally, during the detection period (t_2), the free induction decay is recorded. During a 2D pulse sequence the evolution time (t_1) is extended for each point in the indirectly detected dimension. The result is a two dimensional data set that is recorded in time consisting of the evolution time and detection time, $s\{t_1,t_2\}$. Fourier transformation of each time domain, proceeding first in the directly detected dimension t_2 and then in the indirectly detected dimension t_1, yields the 2D NMR spectrum, $S\{\nu_1,\nu_2\}$.

2D NMR and Peptide Structure Determination

The strategy for structure determination has been outlined by K. Wüthrich in his book, *NMR of Proteins and Nucleic Acids* (6). Peptide structure determination by generally starts by an assignment process such that amino acid spin systems are identified from a 2D spectrum. Experiments used for the assignment process

utilize the scalar coupling network such that intraresidue protons that are scalar coupled can be correlated by their chemical shift. TOCSY (12) spectra are especially suited to meet this end since complete amino acid spin systems are obtained from this experiment. The TOCSY pulse sequence is given in Fig 1-3. The TOCSY experiment, also known as the HOHAHA experiment (12), utilizes a mixing time composed of several composite pulses. The mixing time serves to remove the Zeeman term from the Hamiltonian such that magnetization is transferred down a scalar coupled network via the strong coupling Hamiltonian. The result is a spectrum where all chemical shifts are correlated for all resonances of a scalar coupled network, provided the mixing time is long enough. The results of the TOCSY spectrum are a benefit for peptide assignments as whole spin systems can be identified quickly from one spectrum. Sequential assignments cannot be made at this point, but spins systems are identified by amino acid type. Where several assignment possibilities exist, a list of amino acid candidates are given to that spin system.

The resonance assignments made according to the TOCSY are supplemented with COSY (13) spectra to confirm the intraresidue proton assignments. Figure 1-4 shows the pulse sequence for the gradient enhanced COSY experiment or GCOSY (14). COSY spectra, similar to TOCSY, correlate the chemical shifts of protons that are scalar coupled. If two protons are scalar coupled a cross peak will be present in the 2D NMR spectrum at the resonance frequencies of the two coupled protons. The gradient enhanced version uses pulsed field gradients to select the desired coherence path for detection, thus

eliminating the need to use a multi-step phase cycle to achieve the same outcome, reducing the overall experimental time and eliminating spectral artifacts.

The amino acid sequence is known *a priori* and the resonance assignments are then mapped directly onto the peptide sequence using a through space technique, specifically NOESY (15) spectra are used to meet this end. The NOESY pulse sequence is shown in Fig. 1-5. During the NOESY mixing time, cross relaxation between spins that are close in space leads to incoherent transfer of magnetization, resulting in the build up of an NOE between the nuclei of interest. The intensity of the resulting cross peak observed in NOESY spectra is proportional to the cross relaxation rate between the nuclei of interest and hence the internuclear distance (6), shown below,

[10] $\sigma_{ab} \alpha 1/r_{ab}^{6}$

where σ_{ab} is the cross relaxation rate between nuclei a and b and r is the distance between the nuclei. The distance dependence of the cross peak is an important result since it gives a direct method to determine important structural information. The r^{-6} distance dependence also indicates that the effect disappears quickly as distances increase and, in fact, the upper bound distance is only 5Å (6). However, the short distance dependence of the cross peak is not entirely a limitation, since protons of the molecule that are far apart sequentially can be found close in space, and will thus define the tertiary structure of the molecule. Of course, the NOE is also essential for sequentially assigning the resonances of the peptide of interest.

The sequential assignment procedure generally starts from a unique portion of the amino acid sequence (i.e. Pro or Arg amino acids make a good starting

point since their spin systems are unambiguous as long as they can be wholly identifiable from TOCSY or COSY spectra). Through space sequential connectivities of the type $d_{\alpha N}$, $d_{\beta N}$, and d_{NN} are built among the assigned resonance signals and then mapped to the amino acid sequence of the macromolecule until all sequential assignments have been made. The NOESY spectra are then examined to find areas of secondary structure. Regular patterns of NOE connnectivities have been outlined, and Fig. 1-6 indicates the expected NOE pattern for the various types of secondary structure (6).

Chemical Shift

Chemical shifts have been used as a method for structure determination. Specifically, the chemical shift index ($\Delta\delta=\delta_{expt}-\delta_{random\ coil}$) (16) has been used as a preliminary step for structure determination of proteins and peptides. It has been noted (14) that for the H_α protons belonging to amino acids in alpha helices there is an average up-field shift of 0.39ppm, and for amino acids in beta sheets there is an average downfield shift of the H_α protons of 0.39ppm. Local areas of secondary structure can be identified quite quickly by identifying a series of consecutively shifted H_α protons. Generally, if four or more amino acids show a deviation from the random coil value greater or less than 0.1ppm (17) this indicates an area of beta sheets or alpha helices, respectively. Nearest neighbour effects should be taken into consideration when applying the chemical shift index (18), coupled with proper referencing of spectra to determine accurate chemical shifts (19).

Although the chemical shift index provides a fast method for determining local areas of secondary structure, other experimental evidence is often required, such as NOE and scalar coupling information, for secondary and tertiary structure determination.

Experimental distances

Molecular modeling involves collecting experimental distances from NOESY spectra, and it is particularly attractive to obtain structural information from a single NOESY spectrum. Consequently, the choice of the experimental parameters for a NOESY experiment is very important. The NOESY spectrum should be collected in the initial rate regime, since the peak intensities will be proportional to the cross relaxation rate and internuclear distance, hence short mixing times are required. Distances can then be estimated from the peak intensity by calibration of the cross peak intensity of interest ($I_{interest}$) with the intensity of a group with a known distance (I_{ref}), shown below.

[11] $r_{interest} = r_{ref}(I_{ref}/I_{int})^{1/6}$

In this equation, $r_{interest}$ is the distance between the nuclei of interest, and r_{ref} is the reference distance. Equation 11 is known as the isolated spin-pair approximation. It must be noted that the distances obtained from NOESY spectra are not precise, as spin diffusion and uncorrelated motions within a peptide may impair the accuracy of the derived distances. Therefore, distances are generally given as bounds as they incorporate errors that are inherent in the distance estimates. Three distance bounds are given and distances are classified accordingly. The distance bounds are strong, medium and weak, and upper bound distances of 2.7Å, 3.4Å,

and 5.0Å are given, respectively. A lower bound distance of 1.8Å, or the sum of two van der Waals radii for two hydrogen atoms is also given. Methyl, methylene, and other groups that contain more than one proton that cannot be stereo-specifically assigned are applied as a pseudo atom (20). The pseudo atom is placed at the geometric centre of the group of interest and a distance correction for the use of the pseudo atom is applied to the upper bound distance. The distance correction is the distance from the pseudo atom to the protons that it replaces.

Molecular Modeling

Several methods have been proposed as a means to model the tertiary structure of a macromolecule. These include distance geometry, simulated annealing or relaxation matrices, and combinations thereof. In the research presented in this thesis, simulated annealing was used exclusively.

Simulated annealing (21,22) falls into the realm of molecular dynamics. Molecular dynamics follows the motions of atoms in a system by numerically solving Newtons's equations of motion, given below (23),

[12] $dE/d\mathbf{r} = md^2\mathbf{r}/dt^2$

where E is the potential energy function that describes the system, m is the mass of the nucleus, and **r** represents the Cartesian coordinates of the nucleus. The goal of simulated annealing is to minimize a target or potential energy function to find the global minimum on the potential energy surface. NMR data alone cannot sufficiently describe the structure of the macromolecule, thus the potential energy function contains terms that describe the covalent geometry and experimental data

to describe the tertiary structure of the molecule. This is given in Equation 13 (24).

[13] $E = E_{Struct} + E_{Expt}$

The first term in Equation 13 describes the structure of the molecule and is made up of several terms for the covalent geometry, including terms for bonds, angles, dihedral angles, nonbonded interactions, and improper dihedral angles. Improper dihedral angles are used to describe chirality and planarity of groups within the molecule, for example chiral centres and planar aromatic rings, respectively. The full form of the term for local covalent geometry is given by (23,24),

[14] $E_{struct} = k_{bond}(r-r_{eq})^2 + k_{dihed}(1+\cos(\phi+\delta)) + k_{improber}(\phi-\delta)^2 + k_{bond\ angle}(\theta-\theta_{eq})^2 + k_{rep}(\max(0,(sR_{min})^2 - R^2))^2$

where the k's are the energy constants for each term, r is the distance between the nuclei, the subscript eq is the reference value, n is the periodicity, ϕ and θ are the measured angles, δ is the phase shift or offset of the angle measured in degrees. The final term in Equation 14 is the repulsive van der Waals term (19,20,21,22), where s is a scaling factor, R_{min} is the van der Waals radii for the atoms of interest, and R is the actual distance between the atoms. The form of the repulsive van der Waals term is such that following conditions are imposed,

[15] $E_{VDW} = 0$ if $(sR_{min})^2 > R^2$

$E_{VDW} = k_{rep}(\max(0,(sR_{min})^2 - R^2))^2$ if $(sR_{min})^2 < R^2$

indicating that a repulsion is applied to the atoms of interest when they are closer than their van der Waals radii.

The experimental term in the potential energy function has terms for NOE distances and dihedral angles that are determined from scalar coupling constants, given by (23),

[16] $E_{Expt} = k_{NOE}\Delta_N^2 + k_{Dihed}\Delta_D^2$

where the delta's are the deviations from the experimental restraints.

Simulated annealing starts with an extended structure of the peptide of interest that has regular covalent geometry. The temperature of the system is set to a high value, 1000K, and the energy constants of the potential energy function are scaled to a low value. The modified form of the nonbonded interaction allows the close approach of nuclei such that during the hot phase of the protocol the molecule is allowed to be freed from local minima and cross barriers on the potential energy surface. The system is then slowly cooled and the energy constants of the potential energy function are scaled to their full value, trapping the molecule in the global minimum on the potential energy surface.

The final stage of the protocol involves Powell or conjugate gradient minimization to send the molecule further down the minima in which it is trapped. Several cycles are performed during this time applying experimental and covalent geometry restraints. The final structure is then evaluated in terms of convergence of acceptable covalent geometry and experimental restraints.

The generalized modeling strategy is an iterative approach. First, unambiguous experimental data is applied to the system and the molecule is modeled. The model is evaluated for restraint violations, which, if present, are removed and the molecule remodeled. Preliminary models are then used to

resolve experimental ambiguities by comparing the modeling results with spectra and the molecule is then remodeled. Molecular modeling proceeds in this fashion until all of the experimental data is input in the modeling protocol.

Table 1-1: The idealized dihedral angles for various turns found in protein structure (1,25,26).

Turn	ϕ_{I+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}
Type I	-60°	-30°	-90°	0°
Type I'	60°	30°	90°	0°
Type II	-60°	120°	80°	0°
Type II'	60°	-120°	-80°	0°
Type III	-60°	-30°	-60°	-30°
Type III'	60°	30°	60°	30°
Gamma	75°	-69°		
Inverse Gamma	-75°	69°		

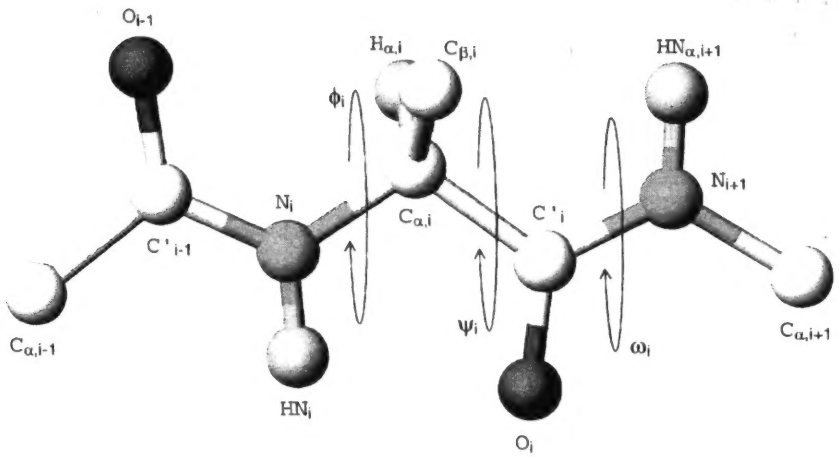


Figure 1-1. The polypeptide backbone showing the defining dihedral angles. Atoms belonging to the amino acid are designated by i, atoms that belong to the preceding amino acid are designated by i-1, and atoms belonging to the following amino acid are designated by i+1. Figure 1-1 was created using the program Molmol (27).

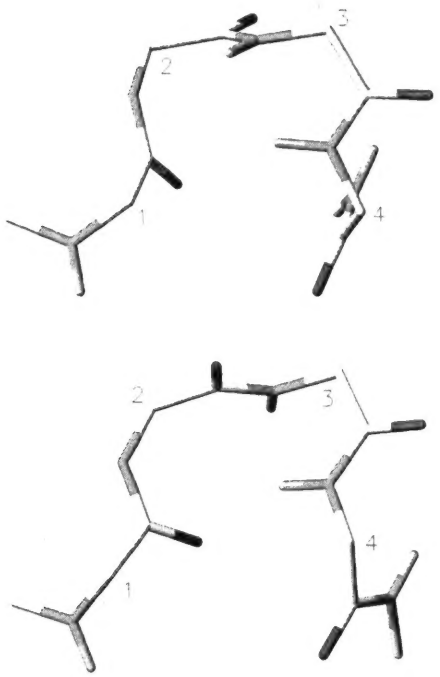


Figure 1-2. Idealized type I and type II beta turns, top and bottom figures, respectively. The positions of the residues within the turn are indicated. Figure 1-2 was created using the program Molmol.

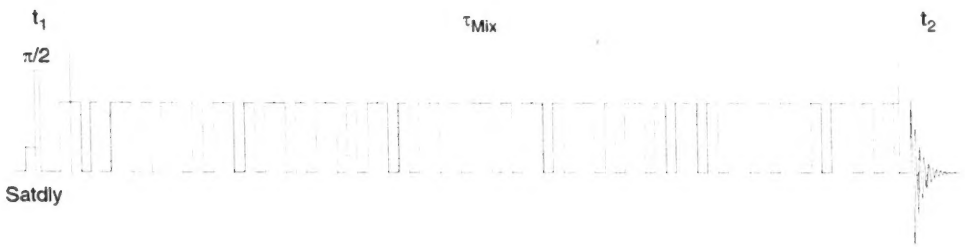


Figure 1-3. The TOCSY pulse sequence with presaturation of the water peak for solvent suppression. Presaturation of the water peak is achieved by the use of a weak continuous radio frequency wave output by the transmitter on resonance with the water peak. Presaturation occurs during the relaxation delay at the start of the pulse sequence and is indicated by the section labeled Satdly. The evolution, mixing and detection periods are indicated. The pulse sequence is taken from VNMR 6.1b.

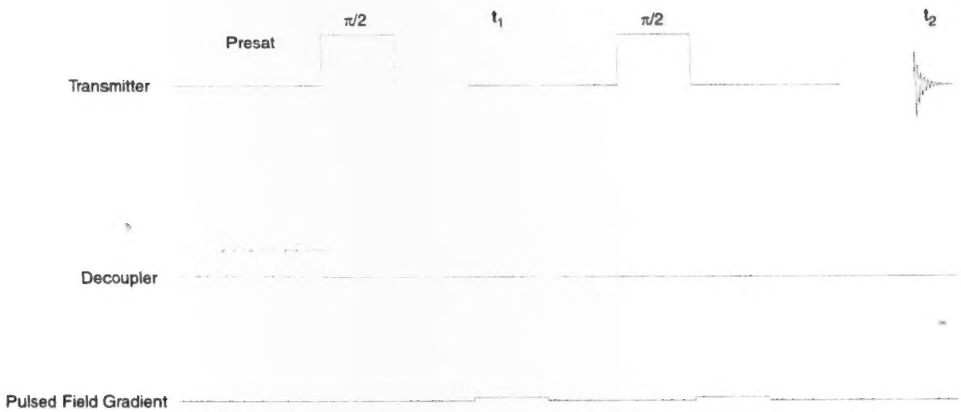


Figure 1-4. The pulse sequence for a GCOSY with presaturation of the water peak for solvent suppression. Presaturation of the water peak is achieved by the use of a weak continuous radio frequency wave, that is output by the decoupler on resonance with the water peak. Presaturation occurs during the relaxation delay at the start of the pulse sequence and is indicated by the section labeled Presat. The evolution and detection periods are indicated. The pulse sequence is taken from VNMR 6.1b.

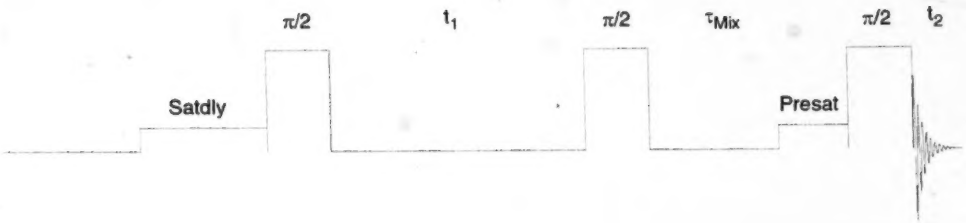


Figure 1-5. The pulse sequence for a NOESY experiment with presaturation of the water peak for solvent suppression. Presaturation of the water peak occurs during the relaxation delay, and is indicated by the section of the sequence labeled Satdly, and during the mixing period indicated by the section labeled Presat. The transmitter is used for solvent suppression during this pulse sequence. The various time periods of the basic 2D experiment are indicated. The pulse sequence is taken from VNMR 6.1b.

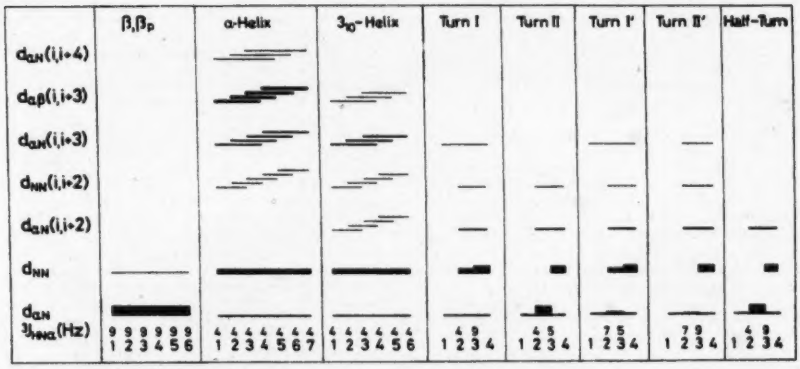


Figure 1-6. The expected NOE's and coupling constants for sections of regular secondary structure (6).

References

1. T.E. Creighton , *Proteins: Structures and Molecular Properties*, 2nd ed., W.H. Freeman: New York, 1993.

2. K. Wüthrich, C. Grathwohl, *FEBS Lett.*, **43**, 337-340, (1974).

3. G.N. Ramachandran and V. Sasisekharan, *Adv. Protein Chem.*, **23**, 283-437 (1968).

4. A.L. Morris, M.W. Macarthur, E.G. Hutchinson, and J.M. Thorton, *Proteins-Struct. Funct. Genet.*, **12**, 345-364 (1992).

5. C.W. Wilmot and J.M. Thorton, *J. Mol. Biol.*, **203**, 221-232 (1988).

6. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John-Wiley and Sons: New York, 1986.

7. G.M. Clore and A.M. Gronenborn, *CRC Crit. Rev. Biochem. Mol. Biol.*, **24**, 479-564 (1989).

8. M. Karplus, *J. Phys. Chem.*, **30**, 11-15 (1959).

9. J. Cavanagh, W.J. Fairbrother, A.G. Palmer III, and N.J. Skelton, *Protein NMR Spectroscopy. Principles and Practice.*, Academic Press: San Diego, 1996.

10. J. Jeener, *Amphère Summer School,* Basko Polje, Yugoslavia, 1971.

11. O.W. Sørenson, G.W. Eich, M.H. Levitt, G. Bodenhausen, and R.R. Ernst, *Prog. NMR Spectros.*, **16**, 163-1692 (1983).

12. A. Bax and D.G. Davis, *J. Magn. Reson.*, **65**, 355-360 (1985).

13. W.P. Aue, E. Barholdi, and R.R. Ernst, *J. Chem. Phys.*, **64**, 2229-2246 (1975).

14. R.E. Hurd, *J. Magn. Reson.*, **87**, 422-428 (1990).

15. J. Jeener, B.H. Meier, P. Bachmann, and R.R. Ernst, *J. Chem. Phys.*, **71**, 4546-4546 (1979).

16. D.S. Wishart, B.D. Sykes, and F.M. Richards, *J. Mol. Biol.*, **222**, 311-333 (1991).

17. D.S. Wishart, B.D. Sykes, and F.M. Richards, *Biochemistry*, **31**, 1647-1651 (1992).

18. D.S. Wishart, C.G. Bigam, A. Holm, R.S. Hodges, and B.D. Sykes, *J. Biomol. NMR*, **5**, 67-81 (1995).

19. D.S. Wishart and D.A. Case, *Methods Enzymol.*, **338**, 3-35 (2001).

20. K. Wüthrich, M. Billeter, and W. Braun, *J. Mol. Biol.*, **169**, 949-961 (1983).

21. M. Nilges, A.M. Gronenborn, A.T. Brünger, and G.M. Clore, *Protein Eng.*, **2**, 27-38 (1988).

22. M. Nilges, G.M. Clore, and A.M. Gronenborn, *FEBS Lett.*, **239**, 129-136 (1988).

23. P. Güntert, *Q. Rev. Biophys.*, **31**, 145-237 (1998).

24. A.T. Brünger, *X-PLOR, Version 3.1: A system for X-ray crystallography and NMR*, Yale University Press: New Haven, 1993.

25. G.D. Rose, L. Gierasch, L., and J.A. Smith, *Adv. Protein. Chem.*, **37**, 1-109, (1985).

26. E.J. Milnerwhite, E.J.; B.M. Ross, B.M.;R. Ismail, K. Belhadjmostefa, and R. Poet R. *J. Mol. Biol.*,**204**, 777-782 (1988).

27. R. Koradi, M. Billeter, and K. Wüthrich, *J. Mol. Graph.*, **14**, 51-55 (1996).

Chapter 2.

Correlation of Secondary Structures of Bradykinin B1 Receptor Antagonists with Their Activity.[1]

[1] A version of this chapter has been published.

M. Miskolzie, L. Gera, J.M. Stewart and G. Kotovych, *J. Biomolec. Struct. Dyn.*, **19**, 585-593 (2002).

Introduction

The goal of this research is the determination of medicinally important peptide conformations in order to understand their biologically active conformations. A model for the peptides bound to their specific receptors is actually required but these receptors have only been isolated in very small amounts and are not available for binding studies. These data will be valuable in the design of newer, more specific drugs and in understanding the mode of action of BK antagonists. Bradykinin (BK) (Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-Pro^7-Phe^8-Arg^9) is a naturally occurring and potent pro-inflammatory peptide whose generation in tissues and body fluids elicits many physiological responses including vasodilation, smooth muscle spasm and edema, as well as pain and hyperalgesia. The biological responses to BK are mediated by two different receptors, called B_1 and B_2. B_2 receptors are constitutively expressed on many cells and mediate all of the normal physiological responses to BK, as well as acute pathophysiological responses. The B_2 receptors require the full chain of the BK molecule for agonist action. In chronic inflammation, a new class of BK receptors, called B_1, is expressed (1). These do not respond to intact BK, but only to des-[Arg^9]-BK. These receptors are typical G-protein-coupled receptors, and have seven transmembrane helices (1). In the absence of purified receptor, we have evaluated the BK antagonist conformations in solvents such as 60% CD_3OH/40% H_2O, or solutions of sodium dodecylsulfate micelles to correlate their secondary structures with their activities. We have completed an NMR study of B-9340 - a third generation antagonist active against both the B_1 and B_2

receptors (2). The sequence is: B-9340: DArg^0-Arg^1-Pro^2-Hyp^3-Gly^4-Thi^5-Ser^6-DIgl^7-Oic^8-Arg^9. Hyp is *trans*-4-hydroxy-L-proline, Igl is (α-(2-indanyl)glycine), Oic is (3aS, 7aS) – octahydroindole-2-carboxylic acid and Thi is beta-(2-thienyl)-L-alanine. In B-9340 a non-ideal type II β-turn was observed from Pro^2 to Thi^5 and a type II' β-turn comprising residues Ser^6 to Arg^9. A salt bridge was observed between the Arg^1 side chain and the C-terminal COO^- (on Arg^9). These conclusions have been used by the Williams group at Berkeley very recently (3) to model the arginine dimer and BK by density functional theory to show that stable gas-phase salt bridges exist for BK together with two β-turns, as we observed. A review of BK antagonist conformations has been published (4) and lists a wide range of predominantly B_2 antagonist conformations, together with the proposed models for binding with the receptors.

On the other hand, B_1 receptor antagonist conformations have not been extensively studied. In view of the recent paper on non-competitive pharmacological antagonism at the rabbit B_1 receptor (5) where the effects of six B_1 receptor antagonists were studied, we felt it was important to correlate the secondary structures of a number of these with their measured activities. Specifically, activities were compared for peptides with and without an Igl residue (Fig. 2-1), first introduced in 1996 (6). We therefore chose three peptides that have, respectively, two, one, and zero Igl residues. The first B_1-specific receptor antagonist, studied several years ago (7), B-9858 (Lys^{-1}-Lys^0-Arg^1-Pro^2-Hyp^3-Gly^4-Igl^5-Ser^6-DIgl^7-Oic^8, Fig. 2-1), has only a type II β-turn involving residues 2-5 and two Igl residues.

The second is B-10148 (Lys^{-1}-Lys^{0}-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-Igl^{5}-Ser^{6}-DF5F^{7}-Oic^{8}, where F5F is 2,3,4,5,6-pentafluorophenylalanine, Fig 2-1) (8), which also displays a type II β-turn from Pro^{2} to Igl^{5}. Both peptides exhibited a salt bridge between the guanidino group of Arg^{1} and the carboxylate group of Oic^{8}. Both peptides also exhibited some β-turn character at the C-terminal end despite lacking the C-terminal Arg. This pseudo β-turn structure may aid the B_2 receptor antagonist effects of these peptides.

Activity measurements on six BK B_1 receptor antagonists by Larrivée *et al.* (5), including four peptides containing Igl, showed that the order of potency is B-9858 > B-10148. Both are very potent and insurmountable B_1 receptor antagonists and D-Igl^{7} confers the high potency for B-9858. The Igl^{5} residue makes both peptides non-competitive and non-equilibrium antagonists of the kinin B_1 receptor. This residue is absent in Ac-Lys-[Leu^{8}]des-Arg^{9}-BK, which is a conventional antagonist, and is surmountable, as well as B-9958 (Lys^{-1}-Lys^{0}-Arg^{1}-Pro^{2}-Hyp^{3}-Gly^{4}-CpG^{5}-Ser^{6}-DTic^{7}-CpG^{8}, where CpG is cyclopentylglycine).

To complete the series, in the present paper a B_1-receptor antagonist secondary structure has been determined for a peptide without an Igl residue, since such a receptor antagonist has a completely different behavior in comparison with the Igl-containing peptides. Specifically, we chose B-10324, Fig. 2-1, the 2,3,4,5,6-pentafluorocinnamoyl, (F5C), derivative of B-9958. It is expected to exhibit B_1 receptor antagonist activity similar to B-9958 as both peptides have the same sequence except B-10324 has the F5C protecting group. The conventional B_1-antagonist Ac-Lys-[Leu^{8}]des-Arg^{9}-BK also contains an

N-terminal protecting group and this N-acetylation confers a complete resistance to degradation in serum, relative to the more potent but fragile sequence Lys-[Leu^8]des-Arg^9-BK (9). A complete secondary structure determination of B-10324 thus allows a comparison of secondary structures of B_1-antagonists containing two, one, and zero Igl residues with their activities that have been published.

The choice of solvent for the NMR studies is important and is 60% CD_3OH/40% H_2O by volume, the same solvent system that was used to study B-9858 and B-10148. This structure-inducing solvent system mimics the results of the conformational studies of other BK antagonists in solutions of sodium dodecylsulphate (SDS) micelles (ref. 4, Table I). It is important to use a solvent system that mimics biological membranes because of the proposed mechanistic principle of peptide recognition, namely the lipids of the target cell membrane exert conformational and other constraints on the peptide, enabling it to find its surface receptor and to expedite specific binding (10). This idea has in fact been recently used in the NMR study of the conformation of the pituitary adenylate cyclase activating polypeptide PACAP(1-21)-NH_2, bound to a PACAP specific receptor (11). The authors observed that the C-terminal region (8-21) forms an α-helix as in the micelle-bound PACAP; in fact, the conformational difference between receptor-bound and the micelle-bound states is limited to the N-terminal seven residues (11).

Recently, it was shown that high-affinity binding of peptide agonists to the human B_1 receptor depends on the interaction between Lys^0 of the agonist and the

fourth extracellular domain of the receptor (12). Mutation studies revealed that both the positive charge and the proper spatial orientation of this residue were required for interaction with the B_1 fourth extracellular domain. Consequently, our present study will provide structural information regarding the conformation of the B_1-antagonists, which should prove invaluable in future binding studies of B_1 antagonists with their receptors.

Materials and Methods

NMR Sample Preparation

The BK antagonist sample (6) was 99% pure, as determined by HPLC, and was used directly for all experiments. The sample (5 mg) was dissolved in 700μL 60% CD_3OH /40% H_2O. Initially the peptide was dissolved in 200μL or less than the final volume of water and the pH was adjusted to approximately 6.0 with aliquots of 0.1M NaOH until a final pH of 5.43 was attained. The remaining volumes of water and methanol were added next. Finally, to the standard 5mm NMR tube, 20uL of a 0.1% solution of 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) was added as an internal reference, set to 0.0ppm, to ensure accurate chemical shift measurements. All chemical shifts are quoted relative to the DSS reference.

NMR Spectroscopy

To identify the presence of a hydrogen bond, temperature studies were undertaken. A series of 1D spectra at 5 temperatures, from 5°C to 25°C, were acquired. To achieve accurate temperature readings all temperatures were calibrated against a methanol sample obtained from Varian. The actual

temperatures were 1.5°C, 6.8°C, 12.1°C, 17.4°C, and 22.4°C and the amide temperature coefficients were then determined by linear regression analysis. The temperature study was performed on a Varian Unity 500MHz Spectrometer. Data acquisition and initial Fourier transformations were performed on a Sun Sparc 5/110 computer running VNMR 6.1b and later processed and analyzed on a Sun Ultra 10 workstation also running VNMR 6.1b.

At 500MHz spectral overlap was significant and consequently a higher field was needed. Two-dimensional (2D) total correlation spectroscopy (TOCSY (13)) and nuclear Overhauser spectroscopy (NOESY (14)) experiments were performed on a Varian Inova 800MHz spectrometer at the National High Field Nuclear Magnetic Resonance Centre (NANUC) at the University of Alberta. The TOCSY and NOESY experiments were carried out at 5°C employing mixing times of 100ms and 250 ms, respectively, and the WATERGATE sequence (15) was added for solvent suppression. All 2D spectra were obtained with a spectral width of 9000Hz and 8200Hz in the t2 and t1 dimensions, respectively, and were acquired in the phase sensitive mode using the States-Haberkorn-Ruben method (16). Spectra were taken with 4K data points in t2 with 512 increments in t1 with 32 scans taken per increment. The data were transferred to the Sun Ultra 10 workstation for data processing. Squared sine bell and shifted sine bell functions were added as weighting functions in both dimensions.

The $^{3}J_{\alpha H\text{-}NH}$ vicinal coupling constants were determined from the NH region (7-10 ppm) of the 1D spectrum where spectral overlap was not present.

Molecular Modeling

Structure calculations were carried out on B-10324 using Crystallography and NMR System (CNS, (17)) software. A simulated annealing protocol was used to model B-10324. Cartesian dynamics were utilized for both the hot and cool phases of the protocol in which a total of 6000 steps were employed during the hot phase at 1000K and 3000 steps for the slow cooling phase with a time step set to 0.005ps for both phases. Following the annealing algorithm the calculated structure was subject to 2000 steps of Powell minimization. All starting geometries were from an extended conformation and the initial velocities for all atoms were randomly given. The target function employed for the structural calculations contained terms for covalent geometry, non-bonded contacts in the form of a repulsive van der Waals term, and experimentally derived NOE distances and dihedral angles.

Cross peaks from the NOESY spectrum were integrated and classified as strong, medium, and weak and upper bound distance limits of 2.7Å, 3.3Å, and 5.0Å, respectively, were applied. The diastereotopic β protons of Pro^2 were used as a reference distance (1.78Å) to classify the cross peaks. The lower bound distance limit of 1.78Å, the sum of the van der Waals radii of two hydrogen atoms, was used for all distances. The relevant distance restraints were transferred to a Silicon Graphics $Indigo^2$ Impact 10000 workstation with CNS software for modeling. Pseudo atoms were added for all non-stereospecifically assigned methylene protons and appropriate distance corrections were added (18).

The program Molmol (19) was used for visualization of the calculated structures and to create Fig. 2-3, the 10 lowest energy conformers of B-10324.

Results and Discussion

The CD_3OH/H_2O chemical shift data for B-10324 are given in Table 2-1, and wherever possible, the NH temperature coefficients and $^3J_{\alpha H\text{-}NH}$ coupling constants are also given. The spectra for this peptide exhibited minor peaks from a minor conformer in solution, which can arise from *cis/trans* isomerization of all Xxx-Pro, and Xxx-Pro-like (i.e. Hyp and Tic) peptide bonds. Interestingly, approximately one third of the molecule exists in a minor conformation, based on integration of the peaks in the 1D spectrum. Assignments for the minor conformer, where possible, are given in Table 2-1 in parentheses. We believe that this minor conformer arises from the *cis* peptide bond between Ser^6 and Tic^7. Unfortunately, due to spectral overlap with the water peak (4.97ppm), the methanol OH peak (5.23ppm), the H_α of Ser^6 (4.82) and the H_α of Tic^7 (5.35ppm) and attenuation of these peaks due to solvent suppression, we were unable to identify the exact origins of this conformer, since the diagnostic H_α-H_α(i,i+1) NOE between these residues was obscured. No other H_α-H_α(i,i+1) NOEs were observed for this minor conformer. As a result, only the conformation of the major conformer in solution was considered. The major conformer displayed strong H_α-H_δ(i,i+1) NOEs between all Xxx-Pro and Xxx-Pro-like residues, indicating that these peptide bonds were in the preferred *trans* conformation. Furthermore, an H_α-$H_{1,1'}$(i,i+1) NOE between Ser^6 and $DTic^7$ was observed indicating that the Ser^6-$DTic^7$ peptide bond was in the *trans* conformation.

However, this NOE was partially obscured by the solvent peaks and attenuated by solvent suppression. Secondary evidence for the *trans* conformation was from NOE's that were observed between both beta protons of Ser^6 and $H_{1,1'}$ of $DTic^7$. Finally, HOE-140, another bradykinin antagonist containing a DTic amino acid, has been studied (20) and it was found that DTic had a *trans* peptide bond, as we believe the major conformer is in B-10324.

Spectral assignments were performed according to standard procedures (21). Individual residue spin systems were identified in the TOCSY spectrum and sequential assignments were determined from the αH-NH and NH-NH regions within the NOESY spectra. Furthermore, the sequential assignments (αH-NH and NH-NH) were used to confirm the spin system assignments determined from the TOCSY spectrum.

The spectra for B-10324 indicate the presence of a 2-5 β-turn (7,8). A medium range αH-NH(i,i+2) NOE from Hyp^3 to Cpg^5 was present, as seen from the αH-NH region of the NOESY spectra for B-10324 shown in Figure 2-2 where the important αH-NH(i,i+2) NOE is highlighted. Also, the expected NH-NH(i,i+1) NOEs were present; a more intense NOE between Gly^4 and CpG^5 than between CpG^5 and Ser^6 was observed. The sequential NH-NH(i,i+1) NOE between CpG^5 and Ser^6 may be suggestive of a second turn centered on Gly^4 and CpG^5, however additional evidence for this structural feature was not observed; that is, an αH-NH(i,i+2) NOE between Gly^4 and Ser^6 was not observed that would confirm the presence of a second β-turn in this area. The presence of the 2-

5 β-turn was also supported by NOEs from the ring protons of CpG^5 to the protons of Pro^2 indicating that the peptide forms a 2-5 β-turn.

The NH proton temperature studies used to determine whether the amide protons are sequestered in a hydrogen bond or shielded from the solvent indicated that CpG^5 was indeed involved in hydrogen bonding and shielded from the solvent. The temperature coefficient for CpG^5 in B-10324 was -4.0 ppb/K.

Molecular Modeling

A total of 76 NOE based distance restraints were used to model the structure of B-10324 (41 intraresidue and 35 interresidue restraints). One relevant coupling constant ($^3J_{C'H\text{-}NH}$) was added for CpG^5, to constrain its phi torsion angle, and one hydrogen bonding restraint was added between residues Pro^2 and CpG^5.

A total of 100 structures were generated and optimized in the simulated annealing protocol. The ten lowest energy conformers that satisfied all distance and dihedral constraints (i.e. no distance violations greater than 0.25Å nor dihedral violations greater than 5° for distance and dihedral angle restraints, respectively) are used to represent the solution structure of B-10324. The conformations of the ten lowest energy conformers were examined further, Fig. 2-3. The pair wise RMSD for the family of conformers is 0.39 +/- 0.15Å for all heavy atoms of the backbone within the area of the turn, namely residues one to five, as calculated by Molmol.

Even though a small NH temperature coefficient was observed for this peptide, a hydrogen bond constraint was not initially applied in the molecule modeling, as the appropriate acceptor atoms could not be identified. To identify

the hydrogen bond, all restraints were initially input and possible hydrogen-bonding interactions were then examined. Preliminary results did indicate the presence of a 2-5 hydrogen bond, as expected of a type II β-turn, and the carbonyl oxygen of Pro^2 was identified as the hydrogen bond heavy atom acceptor for the peptide. Thereafter, the hydrogen bond was constrained by enforcing a lower bound distance of 1.4Å and an upper bound distance of 2.7Å between the amide proton of CpG^5 and the hydrogen bond acceptor atom, as well as a constraint of 2.4Å to 3.7Å for the heavy atom donor and acceptor atoms.

The results of the modeling indicate that the only secondary structure element that was observed for B-10324 was a distorted type II β-turn with a hydrogen bond. The average phi/psi angles for the central amino acids in the region of the turn are -52.8^{o} and 161.9^{o}, respectively, for Hyp^3 and 83.8^{o} and -30.5^{o}, respectively, for Gly^4.

In contrast to the peptides B-9858 and B-10148, B-10324 did not show any evidence for a salt bridge between the guandino group of Arg^1 and the carboxylate group at the C-terminus. The diastereotopic delta protons of Arg^1 were equivalent (7,8) in the NMR spectra, Table 2-1. In fact, it is apparent from Fig. 2-3 that both the N and C-termini are random and mostly extended, as there were few structurally important NOEs observed in these regions.

Previous studies (22) on BK B_1 agonists carried out on a des-Arg^9 analogue in the presence of micelles suggested that the selectivity for the B_1-receptor is caused simply by the lack of the positive charge arising from the absent C-terminal arginine residue. Our studies also indicate that the N-terminal 2 to 5 β-

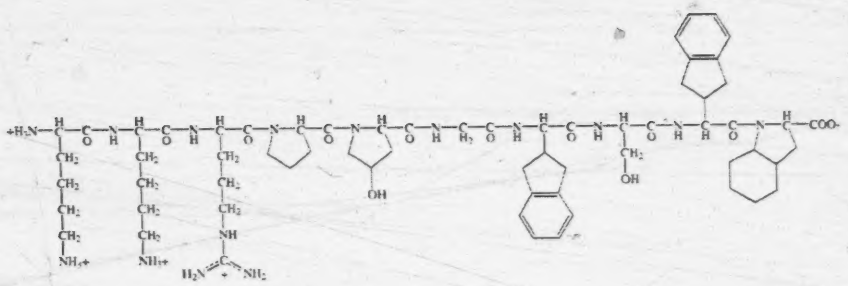
turn is important for antagonist activity. The structure of B-10324, a des-Arg^9 BK B_1 receptor antagonist, is consistent with these findings. In addition, the biological data indicate that the order of potency is B-9858>B-10148 for the Igl^5 containing peptides. Both of these have a type II β-turn involving residues 2 to 5 and residual structure involving residues 6 to 8, even in the absence of the Arg^9 residue. They also have a salt bridge involving the C-terminal carboxylate and the side chain of Arg^1. The biological data indicate that these are non-competitive, nonequilibrium antagonists of the kinin B_1 receptor. By comparison, B-10324, in analogy with B-9958, is a surmountable antagonist. The secondary structure determined in this paper indicates a distorted type II β-turn involving residues 2 to 5, the absence of a salt bridge, possibly due to the presence of the F5C group, and the absence of residual structure at the C-terminal end. These are small differences in the secondary structure and the biological differences between the two types of antagonists therefore rests with the presence or absence of the Igl residue.

Table 2-1: The proton chemical shifts, coupling constants and NH temperature coefficients of B-10324*.

	NH	α,α'	ββ'	γ,γ'	δ,δ'	Other	$^3J_{\alpha H\text{-}NH}$	T.C. (-ppb/K)
F5C	-	7.00	7.49					
Lys^{-1}	8.80 (8.83)	4.42 (4.42)	1.87,1.77	1.49	1.71	2.99	5.95	5.04
Lys^{0}	8.57	4.33	1.82,1.74	1.70	1.48	2.99	5.95	-
Arg^{1}	8.55 (8.52)	4.61 (4.61)	1.85,1.75 (1.82,1.73)	1.70 (1.63)	3.19 (3.16)	ε-NH 7.52 (ε-NH 7.25)	6.77	-
Pro^{2}	-	4.72	2.38,1.94	2.07	3.92,3.52			-
Hyp^{3}	-	4.54	2.33,2.04	4.60	3.85,3.81			-
Gly^{4}	8.71 (8.66)	4.05,3.77 (4.06,3.83)					5.14 (4.65)	5.58
CpG^{5}	7.94 (8.01)	4.22 (4.21)	2.23,1.77,1.68,1.51,1.37,1.23 (2.20,1.74,1.66,1.59,1.50,1.23)				8.66 (7.83)	3.98
Ser^{6}	8.53 (8.62)	5.11 (4.82)	3.83,3.77 (3.83,3.71)				7.57 (6.81)	-
$DTic^{7}$	-	5.13 (5.36)	3.25 (3.44,3.31)			H1,1' 4.84 Aromatic 7.25,7.22 (H1,1' 4.77,4.53)		-
CpG^{8}	7.64 (7.71)	4.00 (4.04)	2.16,1.57,1.54,1.46,1.13,1.04 (2.13,1.47,1.42,1.00,0.83)				7.22 (8.51)	6.62

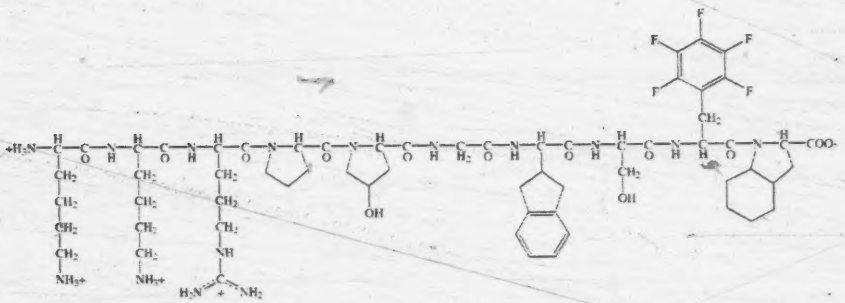
*A minor *cis* conformer was identified in solution. The peaks that could be identified are given in parentheses.

A)



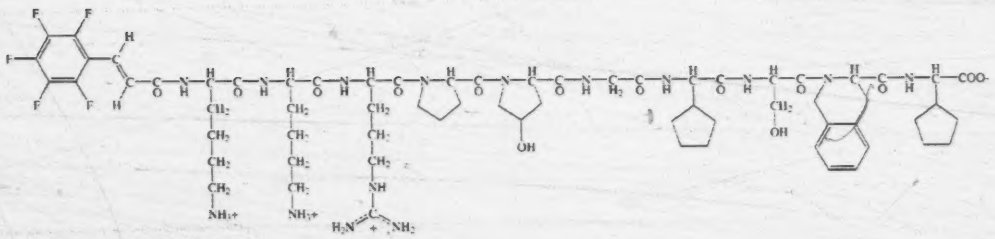
Lys^{-1} Lys^{0} Arg^{1} Pro^{2} Hyp^{3} Gly^{4} Igl^{5} Ser^{6} DIgl^{7} Oic^{8}

B)



Lys^{-1} Lys^{0} Arg^{1} Pro^{2} Hyp^{3} Gly^{4} Igl^{5} Ser^{6} DF5F^{7} Oic^{8}

C)



F5C Lys^{-1} Lys^{0} Arg^{1} Pro^{2} Hyp^{3} Gly^{4} CpG^{5} Ser^{6} DTic^{7} CpG^{8}

Fig. 2-1. The primary structures of the bradykinin B_1 receptor antagonists A) B-9858, B) B-10148, and C) B-10324.

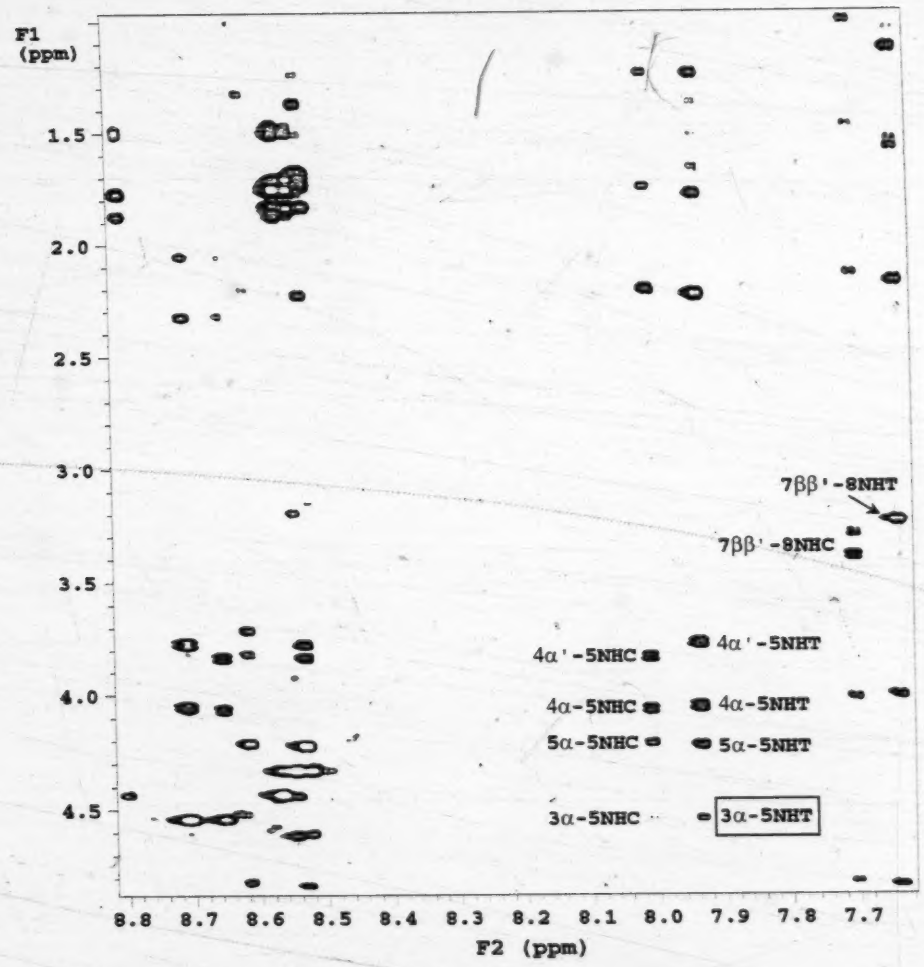


Fig. 2-2. The αH-NH region of the 200ms NOESY at 800MHz and 5°C observed for B-10324. The structurally important NOEs are highlighted for the *trans* (T) conformer and for some of the *cis* (C) conformer.

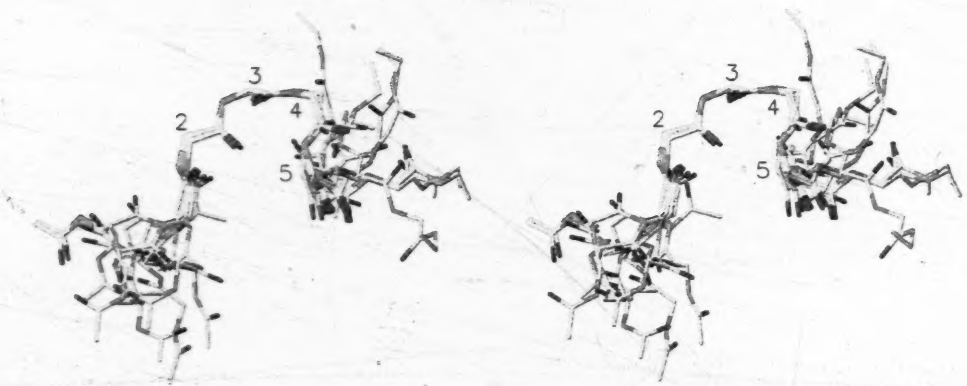


Fig. 2-3. The stereo view of the 10 lowest energy conformers of B-10324 showing the 2-5 type II β-turn. The best fit superposition of the backbone heavy atoms for residues 1-5 is shown.

References

1. F. Marceau, J.F. Hess and D.R. Bachvarov, *Pharmacological Revs.*, **50**, 357-386 (1998).
2. J. Sejbal, J.R. Cann, J.M. Stewart, L. Gera and G. Kotovych, *J. Med. Chem.*, **39**, 1281-1292 (1996).
3. E.F. Strittmatter and E.R. Williams, *J. Phys. Chem.*, **A104**, 6069-6076 (2000).
4. G. Kotovych, J.R. Cann, J.M. Stewart and H. Yamamoto, *Biochem. Cell. Biol.*, **76**, 257-266 (1998).
5. J.-F. Larrivée, L. Gera, S. Houle, J. Bouthillier, D.R. Bachvarov, J.M. Stewart and F. Marceau, *Br. J. Pharmacol.*, **131**, 885-892 (2000).
6. L. Gera, J.M. Stewart, E.T. Whalley, M. Burkard and J.S. Zuzack, *Immunopharmacology*, **33**, 183-185 (1996).
7. J. Sejbal, Y. Wang, J.R. Cann, J.M. Stewart, L. Gera and G. Kotovych, *Biopolymers*, **42**, 521-535 (1997).
8. M. Miskolzie, L. Gera, J.M. Stewart and G. Kotovych, *J. Biomolec. Struct. Dyn.*, **18**, 249-260 (2000).
9. G. Drapeau, R. Audet, L. Levesque, D. Godin and F. Marceau, *J. Pharmacol. Exp. Ther.*, **266**, 192-199 (1993).
10. R. Schwyzer, *Bioploymers*, **37**, 5-16 (1995).
11. H. Inooka, T. Oktaki, O. Kitahara, T. Ikegami, S. Endo, C. Kitada, K. Ogi, H. Onda, M. Fujino, and M. Shirakawa, *Nat. Struct. Biol.*, **8**, 161-165 (2001).
12. D.B. Fathy, D.J. Kyle and L.M.F. Leeb-Lundberg, *Mol. Pharmacol.*, **57**, 171-179 (2000).

13. A. Bax, and D.G. Davis, *J. Magn. Reson.*, **65**, 355-360 (1985).

14. D. Neuhaus, and M.P. Williamson; *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publishers, New York, 1989.

15. M. Piotto, V. Saudek, and V. Sklenar, *J. Biomol. NMR*, **2**, 661-665 (1992).

16. D.J. States, R.A. Haberkorn, and D.J. Ruben, *J.Magn. Reson.*, **48**, 286-292 (1982).

17. A.T. Brünger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, and G.L. Warren , *Acta Cryst.*, **D54**, 905-921 (1998).

18. K. Wüthrich, M. Billeter, and W. Braun, *J. Mol. Biol.*, **169**, 949-961 (1983).

19. R. Koradi, M. Billeter, and K. Wüthrich, *J. Mol. Graph.*, **14**, 51-55, (1996).

20. W. Guba, R. Haessner, G. Breipohl, S. Henke, J. Knolle, V. Santagada, and H. Kessler, *J. Am. Chem. Soc.*, **116**, 7523-7540 (1994).

21. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John-Wiley and Sons, New York, 1986.

22. M. Pellegrini, M. Tancredi, P. Rovero, and D.F. Mierke, *J. Med. Chem.*, **42**, 3369-3377 (1999).

Chapter 3.

The NMR-derived Conformation of Neuropeptide F from *Moniezia expansa.*[1]

[1] A version of this chapter has been published.

M. Miskolzie and G. Kotovych, *J. Biomolec. Struct. Dyn.*, **19**, 991-998 (2002).

Introduction

The neuropeptide Y (NPY) family of peptides consists of NPY, a peptide hormone and neurotransmitter, and the endocrine peptides, peptide YY (PYY) and pancreatic polypeptide (PP) that act at G-protein coupled receptors. Neuropeptide Y is the most prevalent in any tissue (1). The importance of NPY in various physiological responses, such as cardiovascular regulation and the control of food intake, can be seen from the publication of a special issue of Peptides (2) dedicated strictly to NPY research. This family of C-terminally amidated peptides containing 36 amino acids has been extensively studied by NMR, and several representative papers will be discussed.

Bader et al. (3) have very recently studied monomeric porcine NPY bound to dodecylphosphocholine (DPC) micelles. In water solution at 32°C (4) it exists as a dimer, but the dimer interface of NPY is similar to the monomer NPY binding to the micelles. The best-defined α-helical region consisting of residues Asp^{16}-Arg^{33} (Fig. 3-1) is stabilized in the presence of the micelles, and the C-terminal tetrapeptide is structured only in this solvent system. In fact, the α-helical region extends from residues Ala^{14} to Tyr^{36}. The N-terminus consisting of residues Tyr^{1} to Pro^{13} is unstructured. The authors suggest that the binding of NPY to the membrane is a necessary step that precedes receptor binding by preorienting the C-terminal pentapeptide (3). This can be seen from a comparison of the NPY data in aqueous solution (4) with data obtained for NPY bound to DPC micelles (3), which indicates the difference in the relative position of the C-terminal Y^{36} residue in these two solutions. Circular dichroism (CD) studies of a

human NPY mutant in which Q^{34} is replaced by P^{34} (5) has important implications to the present study on NPF, which also has the P^{34} residue. The CD studies on the mutant NPY peptide show that it exists predominantly as a monomer.

The two remaining members of the family will be briefly discussed. Porcine PYY (Fig. 3-1), a monomer at millimolar concentrations in water, consists of two C-terminal helical segments from residues 17 to 22, and 25 to 33 (6), and a turn centered around residues 12 to 14, with the N-terminus folded near residues 30 and 31. The N- and C-termini are brought close together by hydrophobic interactions. Such a folded structure was previously observed in the X-ray study of avian PP, which consists of a polypropylene type helix for residues 1 to 8, a β-turn from residues 9 to 13, and an α-helix originating at residue 14 and terminating at residue 32 (7). There is flexibility in the C-terminal region, for some acidic side chains (D^{15}, E^{26}), and in the β-turn region (8). The overall fold of the bovine PP molecule based on the solution NMR structure (9) was found to be similar to the crystal structure of avian PP. This structural motif has been referred to as the PP-fold (10).

An example of the correlation between solution conformation and receptor subtype selectivity has been shown using CD and NMR studies, together with sedimentation equilibrium analyses (11) on PYY, [P^{34}]PYY, and PYY-(3-36). PYY exhibits nonselective binding with Y_1 and Y_2 receptor subtypes, whereas [P^{34}]PYY is selective for Y_1, and PYY-(3-36) is selective for Y_2. The data show that all three peptides are predominantly monomeric, and the measured helicities are 42%, 31%, and 24% for PYY, [P^{34}]PYY, and PYY-(3-36), respectively. Their

results suggest that small changes in the amino acid sequence may result in large changes in secondary structure and hence in receptor subtype selectivity.

Neuropeptide F differs from the NPY family in that residue 36 is always F^{36} (Fig. 3-1), and the last four amino acids have the RXRF-NH_2 sequence motif. It was originally isolated from the flatworm (platyhelminthes) *Moniezia expansa* (12), a parasitic tapeworm found in sheep intestines. It is the most abundant neuropeptide in platyhelminthes, yet its exact function remains unknown (13). A comparison of NPF with the NPY peptide shows that NPF does not have the conserved $P^2xxP^5xxP^8$ primary structural motif at the N-terminal end and the number of conserved amino acids between these two peptides is five, namely L^{17}, Y^{20}, Y^{27}, R^{33} and R^{35}. Finally, the C-terminal tetrapetide of NPF, $R^{33}P^{34}R^{35}F^{36}$, is identical to the amphibian and reptilian PP (13). The reason for the extensive research into neuropeptides from worms is the belief that these data will yield novel molecules for new antiparasitic drugs.

Because of the importance of NPF and its high presence in flatworms, the present study involves an NMR conformational analysis of NPF. The correlation of structure and function requires a sound characterization of the structural properties of the neuropeptide. This study will allow a comparison between the structures of the NPY family of peptides and NPF, and will help in understanding the conformations of related neuropeptide F peptides isolated from other invertebrates (13,14).

Methods and Materials

NMR Sample Preparation

NPF, at natural isotope abundance, was purchased from Bachem Bioscience (King of Prussia, Pennsylvania, USA) and was 99% pure as determined by HPLC. The NPF sample was used directly for all NMR experiments. The NMR sample (4.4 mg NPF) was dissolved in 700μL of 60% CD_3OH /40% H_2O by volume. Initially the peptide was dissolved in about 200μL of water and the pH was adjusted to approximately 5.0 with aliquots of 0.1M NaOH and 0.1M HCl. The final pH of the solution was 4.85. Then the remaining volumes of water and methanol were added. All chemical shifts are referenced to the residual methanol methyl peak set at 3.30ppm, downfield from TMS.

NMR Spectroscopy

Homonuclear ^1H total correlation spectroscopy (TOCSY) (15,16), nuclear Overhauser effect spectroscopy (NOESY) (17), and double quantum filter correlation spectroscopy (DQF-COSY) (18,19) spectra for NPF were carried out on a Varian Inova 800MHz spectrometer at the National High Field Nuclear Magnetic Resonance Centre (NANUC) at the University of Alberta. Additionally, a [^{13}C,^1H]-heteronuclear single quantum coherence spectroscopy experiment (HSQC) (20) was acquired on a Varian Unity 500MHz spectrometer in the Department of Chemistry at the University of Alberta to verify the Leu and Ile side chain assignments. All spectra were acquired at 298K in the phase sensitive mode of States-Haberkorn-Ruben (21) and processed with VNMR 6.1b on a Sun Ultra 10 workstation.

The ^{1}H NOESY and TOCSY experiments employed a spectral width of 12000Hz in t2 and 10000Hz in t1. For the TOCSY experiment, a total of 2K data points were taken in t2 and 512 increments were taken in t1 with 16 scans taken per increment with a spin lock mixing time of 100ms. The NOESY experiment was taken with 4K data points in t2 with 1024 increments in t1 with 16 scans taken per increment with a mixing time of 150ms. The TOCSY and NOESY experiments employed the Watergate sequence (22) for solvent suppression. The ^{1}H DQF-COSY spectrum employed a spectral width of 12000Hz in both dimensions and attenuation of the water peak was achieved by presaturation. A total of 4K data points in t2 were taken and 512 increments in t1 with 8 scans taken per increment. All homonuclear spectra were processed using sine bell and shifted sine bell weighting functions in both dimensions. A representative NOESY spectrum is given in Fig. 3-2.

The [^{13}C,^{1}H]-HSQC spectrum was measured with 2K data points in t2 and 512 increments in t1 with 32 transients taken per increment. The spectral width was 5000Hz in t2 and 18750Hz in t1. The decoupler was set to the center of the spectrum in the indirectly detected dimension and a ^{13}C resonance frequency of 125.6MHz was used. Attenuation of the water peak was achieved by presaturation. The spectrum was processed with gaussian weighting functions in both dimensions.

After the NOESY and TOCSY spectra were processed with VNMR 6.1b, they were transferred to the program NMRView (23) for analysis and to generate the restraints file for the structure calculations.

Molecular Modeling

To model NPF, a simulated annealing protocol was employed using the program Crystallography and NMR System version 1.0 (CNS) (24) on a SGI Indigo^2 R10000 workstation using torsion angle molecular dynamics (25). The potential energy function employed in the simulated annealing protocol contained a repulsive van der Waals energy term, as well as terms for bonds, angles, impropers, and experimental distance restraints. No electrostatic terms were included in the potential energy function. Starting from an extended structure, the peptide was subjected to one thousand steps of high temperature annealing at 50000K. The initial velocities of all atoms were randomly given and the force constants for the NOE and dihedral restraints were set to values of 150 kcal/molÅ^{-2} and 200 kcal/molrad^{-2}, respectively. The time step was set to 0.015ps. Two slow cooling phases were used. The first stage employed torsion angle dynamics, the second stage used Cartesian dynamics. For the first slow cooling phase the temperature was reduced from 50000K to 2000K in one thousand time steps; each time step was set to 0.015ps. In the second stage, the system was cooled from 2000K to 0K, utilizing three thousand time steps, with each time step set to 0.005ps. The force constants for the dihedral and distance constraints were left unchanged throughout the cooling stages. The repulsive van der Waals term was scaled to 0.1 times its full value during the hot phase of the protocol, to its full value during the first cooling stage and then to 4 times its full value during the final stage of the protocol. Following the last cooling phase, the structure was further regularized by 2000 cycles of Powell minimization.

The program Molmol (26) was used to visualize the calculated structures and to produce Fig. 3-4, which shows the 20 lowest energy conformers of NPF.

Results and Discussion

The NPY family of peptides has been found to aggregate, usually as dimers, at concentrations suitable for NMR studies (3); hence a dilution study was undertaken to assess the aggregation state of NPF. A fresh NMR sample was prepared using 0.7mg of NPF prepared in a similar manner to that described above. The final pH of the solution was adjusted to 4.85 with aliquots of 0.1M NaOH and 0.1M HCl. A 1D experiment was taken to examine the aggregation state of the peptide. The spectrum was compared to the 1D spectrum of the original sample and no differences were found. Hence, NPF does not undergo self-association in solution. In addition, NOESY cross peaks were not observed that could be interpreted in terms of aggregation.

Sequential resonance assignments were performed using a combination of TOCSY and NOESY spectra, employing mixing time of 100ms and 150ms, respectively, and using standard methods (27). Complete proton-proton sequential assignments, $\alpha N(i,i+1)$ and $NN(i,i+1)$, were not possible along the length of the molecule due to overlap of some resonance signals. Also, coupling constants were not available from the DQF-COSY spectra due to broadening of the peaks. As a result, stereospecific assignments of the β-methylene protons and $^{3}J_{\alpha H\text{-}NH}$ coupling constants could not be determined or included in the structural calculations.

All Xxx-Pro peptide bonds were found to be in the preferred *trans*-conformations as seen by strong $\alpha\delta(i,i+1)$ NOE's. In addition, the amide protons

of Asn and Gln were assigned with the aid of intraresidue NOE's. The ^1H chemical shift data for NPF are given in Table 3-1 and have been deposited in the BioMagResBank (access code 5192). A summary of the NOE data is given in Fig. 3-3.

It is evident that the main element of secondary structure is an alpha helix from residues Lys^{14} to Ile^{31}, as is apparent from the regular pattern of $\alpha N(i,i+3)$ and $\alpha\beta(i,i+3)$ contacts within this region of the molecule (Fig. 3-2 and Fig. 3-3). The alpha helix is also supported by the chemical shift index (Fig. 3-3). It is evident from the NOE contacts (Fig. 3-3) that there is little structure at the N-terminus of the molecule. Residues Pro^{-2} to Asn^{13} display mostly sequential connectivities, indicating that this region of the molecule is unstructured. Similarly, the C-terminus is in a random conformation, as only sequential NOE connectivities are observed for the C-terminal pentapeptide of NPF, with no long range NOE's to these amino acids that would indicate a stable secondary structure.

Restraints for the modeling were determined from the 150ms mixing time NOESY. The program NMRView was used to calibrate the NOE derived distance restraints into three categories, strong, medium, and weak with upper bound distances of 2.7Å, 3.4Å, and 5.0 Å, respectively. A lower bound distance of 1.8Å was applied to all distance restraints. Psuedo atoms were added for all nonstereospecifically assigned methylene protons and the appropriate distance correction added (28). A total of 254 NOE-derived distance restraints (89 intraresidue, 96 sequential, 68 medium range and 1 long range) were used to

model the conformation of NPF. A total of 100 structures were generated in the simulated annealing protocol. The 20 lowest energy conformers that satisfied all distance restraints, with no violations greater than 0.5Å, with regular covalent structure, are used to represent the solution structure of NPF, shown in Fig. 3-4. The structures for NPF have been deposited in the Protein Data Bank under the access code 1K8V.

The 20 lowest energy conformers were examined further. The RMSD of all backbone heavy atoms for the ensemble of lowest energy structures is 6.87 +/- 1.86Å. The RMSD for all heavy atoms for the ensemble is 8.33 +/- 1.73Å. These large numbers indicate the lack of structurally important NOE's at the N-terminus and sidechains of the molecule, and hence conformational diversity at the N-terminus. Within the area of well defined structure, residues 14 to 31, the RMSD of all backbone heavy atoms for the ensemble of lowest energy structures is 0.70 +/- 0.25Å and 1.79 +/- 0.25Å for all heavy atoms as calculated by Molmol. A summary of statistics for the structural calculations is given in Table 3-2.

The calculated structure is consistent with the NOE data: a random N-terminus, including residues Pro^{-2} to Asn^{13}, followed by a helical segment from residues Lys^{14} to Ile^{31}. Gly^{32}, a known helix breaker, terminates the helical segment and the final four C-terminal residues are disordered. Gly^{32} is not involved in the α-helix and is also in a random conformation. The lack of precision seen at the two termini of the molecule is the result of a paucity of NOE information in these areas. These data suggest that there is some flexibility in these regions (Fig. 3-4).

The helix is amphipathic in nature (Fig. 3-5) and is slightly bent along its length with the hydrophobic residues folded towards the concave side of the helix. The amphipathic helix is also supported by several side chain NOE's between hydrophobic residues in the (i,i+3) and (i,i+4) positions along the hydrophobic side of the helix. Specifically, NOE contacts were observed between the side chains of Leu^{17} and Tyr^{20}, Ile^{24} and Tyr^{20}, Ile^{24} and Tyr^{27}, Ile^{31} and Tyr^{27}, and Ile^{31} and Phe^{28} (data not shown). These hydrophobic interactions also serve to stabilize the helix in solution (29, and references therein).

It has been noted that the NPF's from several sources (30) share the C-terminal sequence motif with the NPY family of peptides that includes NPY, PYY, and PP. The C-terminal sequence motif is $[F,Y]^{20}$-$X^{21,22,23}$-$[L,I,V,M]^{24}$-$X^{25,26}$-Y^{27}-$X^{28,29,30}$-$[L,I,V,M,F,Y]^{31}$-X^{32}-R^{33}-X^{34}-R^{35}-$[F,Y]^{36}$-NH_2, where X is any amino acid, the amino acids in brackets are variable, and NH_2 is the C-terminal amide (30). As a result, a comparison of the NPF and NPY structures is desirable. The global structure of NPF is very similar to the structure of monomeric NPY when bound to DPC micelles (3) for which there is a C-terminal α-helix preceded by a random N-terminus. The starting point of the helix is the same in both molecules, residue 14, but the extent of the alpha helix in NPF is shorter than observed in porcine NPY. The alpha helix in porcine NPY extends to the end of the C-terminus, whereas the alpha helix in NPF ends at residue 31. Finally, the nature of the alpha helices in these peptides is similar, as they are amphipathic and slightly bent.

An amphipathic alpha helix is typical of a peptide that interacts at a membrane surface. NPY is known to interact at the membrane surface through the hydrophobic interactions along of the alpha helix. Then the peptide is thought to undergo two-dimensional diffusion to the receptor, where receptor recognition and binding occurs (3). The similarities between these two molecules suggest that NPF may undergo a similar mechanism for receptor binding and recognition.

Invertebrate peptides have become important recently as new and novel sources for drugs, and elucidation of their solution structures will aid in understanding their function (31). Knowledge of the neuropeptide F solution structure will lead to a better understanding of its function and may lead to the design and synthesis of new antiparasitic drugs.

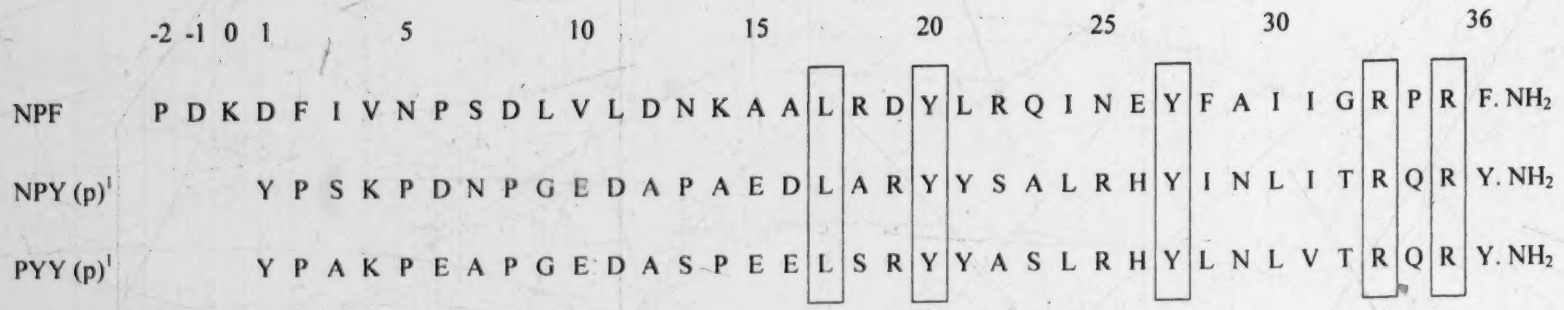
Table 3-1: The ^1H chemical shifts of NPF in CD_3OH/H_2O 60%/40% by volume, referenced to CHD_2OH.

Amino acid	NH	Hα	Hβ(Hβ')	Hγ(Hγ')	Hδ(Hδ')	Other
Pro^{-2}	-	4.34	2.45	2.06	3.44(3.52)	
Asp^{-1}	8.80	4.56	2.74(2.65)			
Lys^{0}	8.38	4.21	1.80(1.77)	1.39	1.64	Hε - 2.96
Asp^{1}	8.15	4.52	2.56			
Phe^{2}	7.92	4.55	3.18(2.98)			Aromatics-7.23
Ile^{3}	7.85	4.13	1.84	1.47(1.13)	0.86	
Val^{4}	7.98	4.10	2.02	0.93(0.90)		
Asn^{5}	8.62	4.99	2.96(2.70)		7.76(6.96)	
Pro^{6}	-	4.27	2.30(2.09)	2.00	3.84(3.93)	
Ser^{7}	8.16	4.22	3.89(3.84)			
Asp^{8}	8.02	4.55	2.79(2.76)			
Leu^{9}	7.64	4.22	1.75	1.62	0.91(0.85)	
Val^{10}	7.61	3.96	2.17	0.99(0.94)		
Leu^{11}	7.96	4.26	1.73	1.59	0.91(0.87)	
Asp^{12}	8.19	4.59	2.80(2.75)			
Asn^{13}	8.38	4.60	2.91(2.84)		7.70(6.91)	
Lys^{14}	8.36	4.13	1.98(1.92)	1.47(1.53)	1.73(1.63)	
Ala^{15}	8.12	4.02	1.52			
Ala^{16}	8.13	4.19	1.49			
Leu^{17}	7.94	4.20	1.83(1.75)	1.72	1.00(0.93)	
Arg^{18}	8.12	3.96	2.05(2.00)	1.89(1.55)	3.23	εNH - 7.45
Asp^{19}	8.23	4.53	2.79(2.72)			
Tyr^{20}	8.26	4.27	3.21			Aromatics-7.01,6.68
Leu^{21}	8.40	3.90	1.94	1.48	0.92	
Arg^{22}	8.12	4.07	2.16	1.81(1.61)	3.29(3.18)	εNH – 7.60
Gln^{23}	8.08	4.00	2.24(2.11)	2.57(2.29)		Hε(ε')-7.31(6.73)
Ile^{24}	8.19	3.74	1.80	1.25(1.00), 0.81	0.58	
Asn^{25}	8.15	4.58	2.98(2.77)		7.64(6.90)	
Glu^{26}	8.41	4.04	2.21(2.07)	2.52(2.28)		
Tyr^{27}	8.09	4.10	3.06			Aromatics-6.60,6.54

Phe^{28}	8.26	4.17	3.22(3.10)			Aromatics-7.38,7.31,7.28
Ala^{29}	8.02	4.12	1.54			
Ile^{30}	7.86	3.87	1.95	1.72(1.19), 0.86	0.86	
Ile^{31}	7.74	4.05	1.80	1.20(1.10), 0.70	0.67	
Gly^{32}	7.87	3.93(3.82)				
Arg^{33}	7.67	4.57	1.82(1.72)	1.67(1.65)	3.14	εNH – 7.34
Pro^{34}	-	4.33	2.20(1.77)	2.00(1.95)	3.76(3.58)	
Arg^{35}	8.31	4.20	1.70(1.64)	1.54(1.50)	3.10	εNH – 7.30
Phe^{36}	7.96	4.59	3.13(2.94)			Aromatics-7.24
NH_2	7.54,7.05					

Table 3-2: Structural statistics for the calculated 20 lowest energy conformers of NPF.

E_{TOT} (kcal/mol)	42.53 +/- 2.05
E_{VDW} (kcal/mol)	9.98 +/- 0.98
E_{NOE} (kcal/mol)	11.78 +/- 1.18
Deviation from Ideal Geometry	
Bonds (Å)	0.0022 +/- 0.00009
Angles (°)	0.2967 +/- 0.0052
Impropers (°)	0.1740 +/- 0.0094


-2 -1 0 1 5 10 15 20 25 30 36
NPF P D K D F I V N P S D L V L D N K A A L R D Y L R Q I N E Y F A I I G R P R F. NH2
NPY (p)[1] Y P S K P D N P G E D A P A E D L A R Y Y S A L R H Y I N L I T R Q R Y. NH2
PYY (p)[1] Y P A K P E A P G E D A S P E E L S R Y Y A S L R H Y L N L V T R Q R Y. NH2

1) p-porcine

Fig. 3-1. Comparison of NPF with porcine NPY and PYY sequences. The boxes indicate the conserved amino acid.

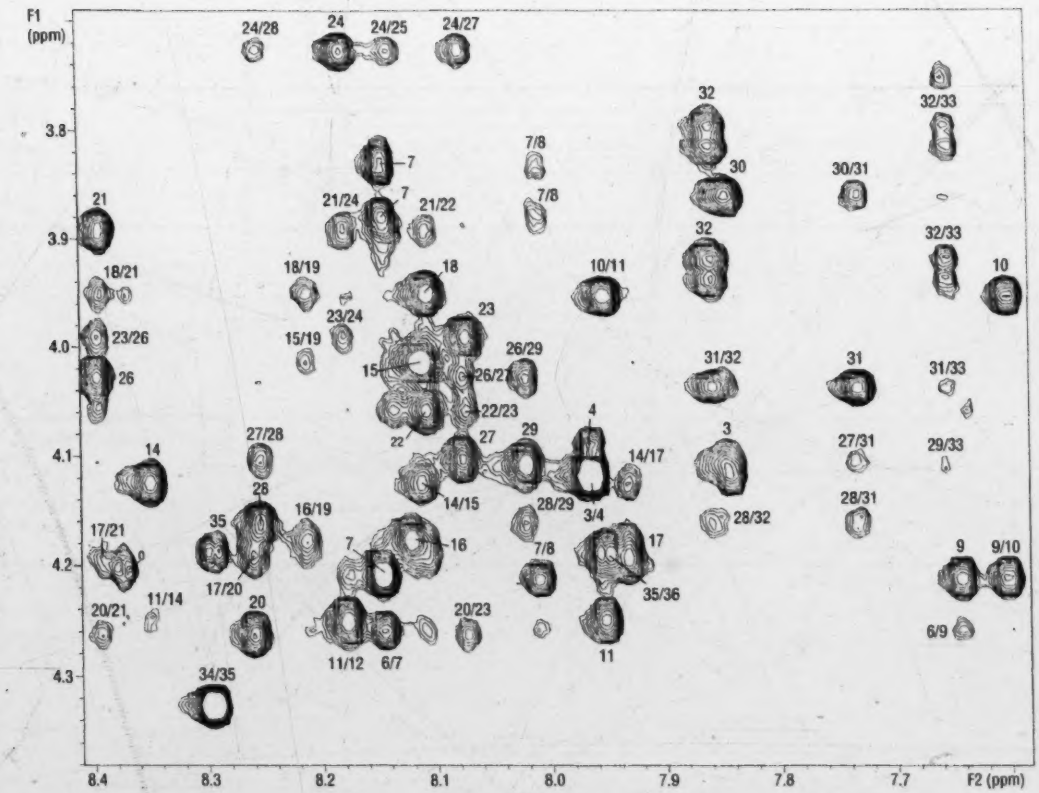


Fig. 3-2. Part of the fingerprint region of the 800MHz, 150ms mixing time, NOESY spectrum for NPF in 60%/40% CD_3OH/H_2O at 298K. The cross peak assignments are given as the sequence position.

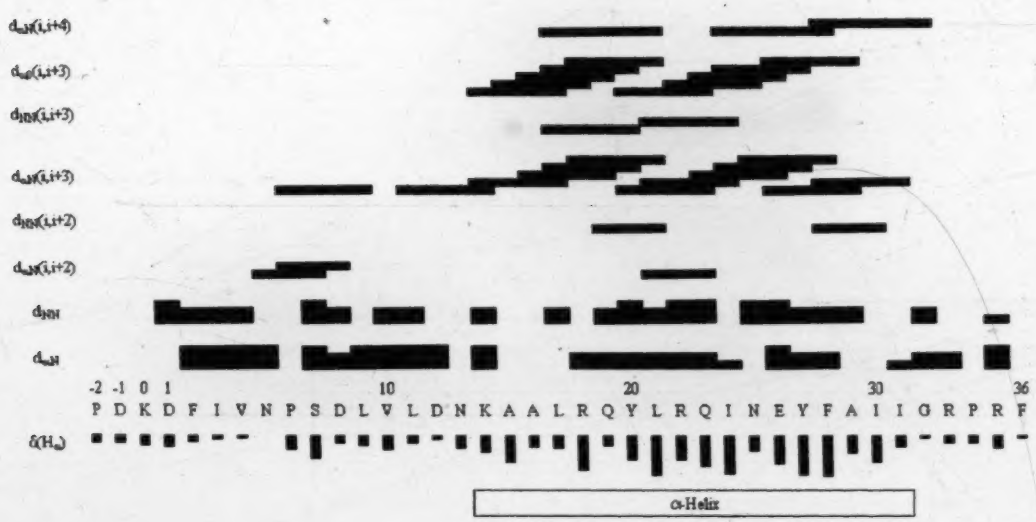


Fig. 3-3. The NOE summary and secondary structure of NPF. The thickness of the bars indicate the intensity of the resonance signals for dαN(i,i+1) and dNN(i,i+1). The medium range connectivities are indicated by bars between the two residues. The chemical shift index is given at the bottom of the Figure.

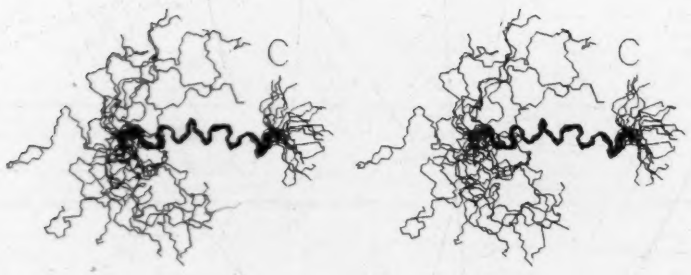


Fig. 3-4. The stereo view of the 20 lowest energy conformers of NPF showing the best fit superposition of the backbone heavy atoms along the helical region, from K^{14} to I^{31}.

A)

B)

Figure 3-5. The twenty lowest energy structures of NPF showing the amphipathic alpha helix. A) The side view of NPF where the C-terminus is indicated, and B) an end on view of NPF with the C-terminus coming out of the page. Hydrophobic side chains are shown in red, hydrophilic side chains are shown in green, and the backbone is shown in blue.

References

1. J.R. Cooper, F.E. Bloom, and R.H. Roth, *The Biochemical Basis of Neuropharmacology*, 7th ed., Oxford University Press, 1996.
2. S.P. Kalra, editor, *Peptides*, **22**(3), Elsevier Press (2001).
3. R. Bader, A. Bettio, A.G. Beck-Sickinger, and O. Zerbe, *J. Mol. Biol.*, **305**, 307-329 (2001).
4. S.A. Monks, G. Karagianis, G.J. Howlett, and R.S. Norton, *J. Biomol. NMR*, **8**, 379-390 (1996).
5. A. Nordmann, M.J.J. Blommers, H. Fretz, T. Arvinte, and A.F. Drake, *Eur. J. Biochem.*, **261**, 216-226 (1999).
6. D.A. Keire, M. Kobayashi, T.E. Solomon, and J.R., Reeve, Jr., *Biochemistry*, **39**, 9935-9942 (2000).
7. T.L. Blundell, J.E. Pitts, I.J. Tickle, S.P. Wood, and C.-W. Wu, *Proc. Natl. Acad. Sci. USA*, **78**, 4175-4179 (1981).
8. I. Glover, I. Haneef, J.E. Pitts, S.P. Wood, D. Moss, I.J. Tickle, and T.L. Blundell, *Biopolymers*, **22**, 293-304 (1983).
9. X. Li, M.J. Sutcliffe, T.W. Schwartz, and C.M. Dobson, *Biochemistry*, **31**, 1245-1253 (1992).
10. J. Fuhlendorff, N.L. Johansen, S.G. Melberg, H. Thogersen, and T.W. Schwartz, *J. Biol. Chem.*, **265**, 11706-11712 (1990).
11. D.A. Keire, P. Mannon, M. Kobayashi, J.H. Walsh, T.E. Solomon, and J.R. Reeve, Jr. *Am. J. Physiol. Gastrointest. Liver Physiol.*, **279**, G126-G131 (2000).

12. A.G. Maule, C. Shaw, D.W. Halton, L. Thim, C.F. Johnston, I. Fairweather, and K.D. Buchanan, *Parasitology*, **102**, 309-316 (1991).

13. T.A. Day, and A.G. Maule, *Peptides*, **20**, 999-1019 (1999).

14. M. De Jong-Brink, A. ter Maat, and C.P. Tensen, *Peptides*, **22**, 309-315 (2001).

15. A. Bax, and D.G. Davis, *J. Magn. Res.*, **65**, 355 (1985).

16. M. Levitt, R. Freeman, and T. Frenkiel, *J. Magn. Res.*, **47**, 328 (1982).

17. S. Macura, and R.R. Ernst, *Mol. Phys.*, **41**, 95-117 (1980).

18. A. Kumar, R.R. Ernst, and K. Wüthrich, *Biochem. Biophys. Res. Commun.*, **95**, 1-6 (1980).

19. M. Rance, O.W. Sorenson, G. Bodenhausen, G. Wagner, R.R. Ernst, and K. Wüthrich, *Biochem. Biophys. Res. Commun.*, **117**, 479-485 (1983).

20. G. Bodenhausen, and D.J. Ruben, *Chem. Phys. Lett.*, **69**, 185-189 (1980).

21. D.J. States, R.A. Haberkorn, and D.J. Ruben, *J. Magn. Reson.*, **48**, 286-292 (1982).

22. M. Piotto, V. Saudek, and V. Sklenar, *J. Biomol. NMR*, **2**, 661-665 (1992).

23. B.A. Johnson, and R.A. Blevins, *J. Biomol. NMR*, **4**, 603-614 (1994).

24. A.T. Brunger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, and G.L. Warren, *Acta Cryst.*, **D54**, 905-921 (1998).

25. E.G Stein, L.M. Rice, and A.T. Brunger, *J. Mag. Res. Ser. B*, **124**, 154-164 (1997).

26. R. Koradi, M. Billeter, and K. Wüthrich, *J. Mol. Graph.*, **14**, 51-55 (1996).

27. K. Wüthrich; *NMR of Proteins and Nucleic Acids*, John-Wiley and Sons, New York, 1986.

28. K. Wüthrich, M. Billeter, and W. Braun, *J. Mol. Biol.*, **169**, 949-961 (1983).

29. S. Padmanabhan, M.A. Jimenez, D, V. Laurents, and M. Rico, *Biochemistry*, **37**, 17318-17330 (1998).

30. C.R. Mair, D.W. Halton, C. Shaw, and A.G. Maule, *Parasitology*, **120**, 71-77 (2000).

31. B.M. Duggan, H.J. Dyson, and P.E. Wright, *Eur. J. Biochem.*, **265**, 539-548, (1999).

Chapter 4.

The NMR-Derived Conformation of Neuropeptide AF, an Orphan G-Protein Coupled Receptor Peptide.[1]

[1] A version of this chapter has been accepted for publication.

M. Miskolzie and George Kotovych, *Biopolymers,* in press (2003).

Introduction

Neuropeptide AF (human NPAF, Ala^1-Gly^2-Glu^3-Gly^4-Leu^5-Asn^6-Ser^7-Gln^8-Phe^9-Trp^{10}-Ser^{11}-Leu^{12}-Ala^{13}-Ala^{14}-Pro^{15}-Gln^{16}-Arg^{17}-Phe^{18}-NH_2) and the closely related peptide, neuropeptide FF (human NPFF, Phe^1-Leu^2-Phe^3-Gln^4-Pro^5-Gln^6-Arg^7-Phe^8-NH_2), belong to a class of peptides known as FMRFamide (Phe-Arg-Met-Phe-NH_2) related peptides. NPFF and NPAF were isolated from the bovine brain using FMRF-NH_2 antiserum (1) and have been identified in other mammalian systems, particularly in humans (2). In fact NPFF and NPAF are widely distributed throughout the mammalian central nervous system and have been indicated as having many different functions (see refs. 3 and 4 for reviews). The most notable functions of these peptides are their anti-opiate and pain modulation effects (1,3, and 4). Interestingly, the gene that encodes for NPFF has been cloned and shown to encode for NPAF and NPFF on the same precursor protein (3,5).

A receptor for these peptides has been identified as a membrane-bound orphan G-protein coupled receptor (6,7), referred to as HLWAR77, for which NPAF and NPFF have a high affinity. Additionally, a second receptor has since been identified and designated NPFF1 (7). The receptor HLWAR77 was also identified in ref. 7 and referred to as NPFF2. Since the anti-opiate NPAF is an endogenous ligand for these receptors, it was suggested (6) that this knowledge will facilitate the identification of the receptor agonists and antagonists dealing with pain modulation and opioid tolerance.

The importance of studying peptides that act at G-protein coupled receptors was illustrated for neuropeptide Y by Bader et al (8). Neuropeptide Y consists of 36 amino acids, is formed in the brain, and the C-terminal amino acids are RQRY-NH_2. It was shown that the binding of neuropeptide Y to the membrane is an essential step preceding receptor binding. This is consistent with the Membrane Compartments Theory proposed by Schwyzer (9,10,11). According to this theory, the target cell membrane induces preferred conformations and orientations of peptides by guiding important residues into the hydrophobic, fixed charge, or aqueous environments of the cell membrane. This has been referred to as "membrane-assisted receptor selection" (8).

In the present study, we determined the conformation of NPAF in trifluoroethanol/water (TFE/H_2O) solutions, and in the presence of the membrane mimetic, sodium dodecylsulphate (SDS) micelles. In addition, the micellar solution was modified with a small amount of the spin label 5-doxyl-stearic acid to investigate the placement of the peptide in the micelle. These structural data may lead to new drug discoveries for the treatment of pain and drug addiction. NPAF was chosen for this study, as the length of the peptide was more amenable for structural studies by NMR.

Materials and Methods

Materials

Neuropeptide AF was synthesized by the Alberta Peptide Institute at the University of Alberta (Edmonton, Alberta). The peptide was purified by HPLC and was determined to be >95% pure by HPLC. The mass of the peptide was

determined by electrospray mass spectrometry. Trifluoroethanol-d_3 (TFE), sodium dodecylsulphate-d_{25} (SDS), and methanol-d_4 were obtained from Cambridge Isotope Laboratories (Andover, MA). 5-doxyl-stearic acid (5-DS) was obtained from Sigma-Aldrich. The peptide and all chemicals were used directly as received for all experiments and sample preparations.

Methods

The NMR Sample

TFE/H_2O Solution

NPAF, 2.9mg, was dissolved in 700μL TFE/H_2O 50%/50% v/v. Initially the peptide was dissolved in 300μL or less than the final volume of water, and the pH was adjusted to approximately 6.0 with aliquots of 0.1M NaOH until a final pH of 6.50 was attained. The remaining volumes of water and TFE were added next. Finally, to the standard 5mm NMR tube, 15μL of a 0.1% w/v solution of 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) was added as an internal reference, and was set to 0.0ppm to ensure accurate chemical shift measurements. All chemical shifts are quoted relative to the DSS reference. The final concentration of the peptide was 2.1mM.

SDS Solution

NPAF, 4.2mg, was dissolved in 700μL H_2O/D_2O, 90%/10% v/v, SDS-d_{25}. Initially, the peptide was dissolved in 500μL or less than the final volume of water, and the pH was adjusted to approximately 4.5 with aliquots of 0.1M NaOH until a final pH of 4.60 was attained. The remaining volumes of water and deuterium oxide were added, and the sample was then added to the required

amount of deuterated SDS. The final peptide concentration was 3.0mM. A total of 53.3mg of SDS-d_{25} was used to make the sample, giving a final SDS concentration of 243mM. The final pH of the sample was 5.65, uncorrected for isotope effects. Finally, to the standard 5mm NMR tube, 15uL of a 0.1% solution of DSS was added as an internal reference.

Spin Label Studies

To examine the placement of NPAF on the micelle, the spin label, 5-doxyl-stearic acid (5-DS) was added to the NPAF SDS solution. A stock solution of 5-DS was made by dissolving 4.5mg of the spin label in 1mL of methanol-d_4; the final concentration of the solution was 0.012M. A 73μL aliquot was removed from the stock solution, added to a vial, and allowed to evaporate to dryness. The NMR sample was then added to the vial to redissolve the spin label (12). The 5-DS was added in a ratio of 1:200 5-DS:SDS, or less than one spin label per micelle. The final concentration of 5-DS was 1.2mM. TOCSY and 1D spectra were taken before and after the addition of the spin label to examine the effects of 5-DS on NPAF.

NMR

Spectra were recorded on Varian Unity 500MHz and Varian Inova 600MHz spectrometers. Data acquisition and initial Fourier transformations were performed on a workstation running VNMR 6.1b, and later transferred to a Sun Ultra 10 workstation also running VNMR 6.1b for processing. One-dimensional (1D) and two-dimensional (2D) total correlation spectroscopy (TOCSY) (13,14), gradient enhanced correlation spectroscopy (GCOSY) (15,16), and nuclear

Overhauser enhancement spectroscopy (NOESY) (17) were carried out. Solvent suppression was achieved by presaturation of the water peak. Spectral widths were equivalent in both dimensions, set to 6000Hz and 7000Hz on the 500MHz and 600MHz spectrometers, respectively. All spectra were acquired in the phase sensitive mode using the States-Haberkorn-Ruben method (18), with the exception of the GCOSY spectrum which was acquired in the absolute value mode selecting N-type signals. The NOESY spectra were acquired with 4K data points in t_2, 512 increments in t_1, and 32 scans were taken per increment. The TOCSY experiment was acquired with 2K data points in t_2 and 256 increments in t_1 with 16 scans per increment. GCOSY spectra were taken with 16 scans per increment with 512 increments in t_1 and 4096 data points in t_2. Squared sine bell and shifted sine bell functions were added as weighting functions in both dimensions. Spectra for the TFE and SDS samples were collected at temperatures of 20°C and 35°C, respectively.

The TOCSY experiments employed a spin lock mixing time of 100ms and 60ms for the TFE/H_2O and SDS samples, respectively, performed on the 600MHz spectrometer. For the TFE/H_2O sample, the mixing time for the NOESY experiment was 150ms, also performed on the 600MHz spectrometer.

For the SDS sample, in an effort to obtain greater sensitivity and remove some of the resonance overlap, a single NOESY experiment was performed on an 800MHz Varian Inova spectrometer at the National High Field Nuclear Magnetic Resonance Center (NANUC) at the University of Alberta. The NOESY spectrum was taken with a mixing time of 150ms and a spectral width of 12000Hz and

10000Hz in the directly and indirectly detected dimensions, respectively. A total of 4K data points were collected in the t_2 and 1024 increments were taken in t_1, with 16 scans per experiment. The temperature was set to 35°C, and attenuation of the water peak was achieved by presaturation.

For the spin label studies, the TOCSY spectra were taken in a similar manner to those described above. However, a total of 4K data points were recorded in t_2 and 512 experiments were recorded in t_1 with 16 scans taken for each experiment. The mixing time was 60ms.

All 2D spectra were processed with VNMR (version 6.1b) and then transferred to the program NMRView (version 5.0.3) (19) for analysis and restraint generation.

Molecular Modeling

A simulated annealing protocol was used to model NPAF (20); the protocol was similar to Brunger's *ab initio* (21) annealing protocol. Structure calculations were started from an extended structure employing r^{-6} sum averaging of distance restraints (22,23). A total of 6000 steps were performed at high temperature (1000K) during which the force constants for the van der Waals terms and dihedral terms were set to a low value. The system was then cooled from 1000K to 0K in 3000 time steps, and the force constants were scaled to their full value. After the cooling stage of the protocol, 2000 steps of Powell minimization were performed to give the peptide regular covalent geometry. The potential energy function employed in the structure calculation used a repulsive van der Waals term, and a term for regular covalent geometry. A term was also included for the

experimental data consisting of NOE based distances and dihedral angles obtained from $^3J_{NH-H\alpha}$ coupling constants. An electrostatic term was not included in the potential energy function.

Structures that had regular covalent geometry and satisfied all of the experimental restraints (that is, no NOE distance violations greater than 0.25Å or dihedral angle violations greater than 5°), and were of lowest energy, were selected to describe the structure of NPAF. The same protocol and acceptance criteria were used for NPAF under both sample conditions.

To model NPAF, the program Crystallography and NMR System (CNS) (24) was used. Restraints for modeling were generated using the program NMRView. Cross peaks were classified as strong, medium, and weak, and upper bound distances of 2.8Å, 3.4Å, and 5.0Å, respectively, were applied. A lower bound distance of 1.8Å was applied to all distance restraints. For amino acids with a $^3J_{HN-H\alpha}$ coupling constant less than 6Hz, indicating that the ϕ dihedral angle is not extended or random, the value of ϕ was set to 60^o+/-30^o.

The calculated structures were visualized using the program Molmol (25) which was also used to prepare Figures 4-4 and 4-8, as well as to calculate the pairwise RMSD for the calculated families of structures.

Results and Discussion

Spectroscopy

Sequential assignments were carried out using standard methods (26). Spin systems were identified in the TOCSY spectra and sequential assignments were made using the NOESY spectra. GCOSY spectra were used to confirm the

intraresidue assignments. Complete sequential assignments, $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$, were not possible due to overlap of some resonance signals, and a combination of $d_{\alpha N}(i,i+1)$, $d_{\beta N}(i,i+1)$, and $d_{NN}(i,i+1)$ were used to make sequential assignments. Also, the side chain amide protons of Asn and Gln amino acids were assigned using intramolecular NOE's. The $^3J_{NH\text{-}H\alpha}$ vicinal coupling constants were determined from the NH region (7-10 ppm) of the 1D spectrum where spectral overlap was not present. The chemical shifts and coupling constants for NPAF in TFE/H_2O and SDS solutions are given in Tables 4-1 and 4-2, respectively. A strong $d_{\alpha\delta}(i,i+1)$ NOE was observed between Ala^{14} and Pro^{15}, indicating that Pro^{15} was in the preferred *trans* conformation in both solutions.

NPAF in TFE

The 1D spectrum of NPAF in TFE displayed several coupling constants less than 6 Hz for residues 7-11 (Table 4-1), indicating that the molecule forms a stable helical conformer in solution. The 2D NOESY spectrum indicated that NPAF was indeed alpha helical, as there were many NOE's diagnostic of a helical structure (26). For example, there were medium range NOE's diagnostic of an alpha helix within the central portion of the molecule (Fig. 4-2) from Leu^5 to Ala^{14} consisting of consecutive $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$, and $d_{NN}(i,i+2)$ NOE's. The αNH region of the NOESY spectrum and the NOE summary chart are presented in Figures 4-1 and 4-2. In addition, the H_α chemical shift index ($\Delta\delta=\delta_{expt}-\delta_{random\ coil}$, Fig. 4-3) (27,28,29) also indicates that NPAF has a helical conformation from Ser^7 to Ser^{11}. This region of the molecule displays the largest up field shifts of the

H_α protons, consistent with the coupling constants that were observed in the 1D spectrum.

On the basis of the above data, NPAF was modeled to determine its tertiary structure in TFE/H_2O. Structural data were obtained from the 150ms mixing time NOESY spectrum taken at 600MHz. To model NPAF, a total of 114 restraints were used, namely 107 NOE based distance restraints (30 intraresidue, 51 sequential, and 26 medium range restraints) and 6 dihedral angle restraints. The dihedral angle restraints were employed to constrain the ϕ dihedral angle of the peptide backbone, based on the $^3J_{NH\text{-}H\alpha}$ coupling constants determined from the 1D spectrum. A total of 200 structures were generated in the molecular modeling protocol. The 20 lowest energy structures that satisfied all distance and dihedral angle restraints, and had regular covalent geometry, were selected to describe the structure of NPAF in TFE and were considered further. The computed structures had no distance restraint violations greater than 0.25Å, or dihedral angle restraint violations greater than 5°. The best fit superposition of residues Asn^6 to Ala^{13}, within the helix for the family of 20 NMR structures, is shown in Figure 4-4A. Figure 4-4B indicates the pairwise RMSD for each amino acid for the best fit superposition of Asn^6 to Ala^{13}. Table 4-3 gives the structural statistics for NPAF in TFE/H_2O.

Residues Asn^6 to Ala^{13} best describe the helical region. The region of the best convergence for NPAF was Ser^7 to Trp^{10}, as all structures converge to an alpha helix in this region with fraying toward either end. The calculated root mean squared deviation (RMSD) for all backbone heavy atoms in this area of the

molecule, Asn^{6} to Ala^{13}, is 0.56+/-0.17; the calculated RMSD for all heavy atoms in this area of the molecule is 1.50+/-0.28, as calculated by Molmol. The N-terminus is disordered due to a paucity of restraints observed in this area of the molecule. This result also suggests that the N-terminus of the molecule is quite flexible.

The NMR spectra suggested that the C-terminal tetra-peptide of the molecule may have some structural features: consecutive $d_{\alpha N}(i,i+2)$ NOE's from Ala^{14} to Gln^{16} and Pro^{15} to Arg^{17}; consecutive $d_{\alpha N}(i,i+3)$ NOE's Ala^{14} to Arg^{17} and Pro^{15} to Phe^{18} (Fig. 4-2) together with strong $d_{NN}(i,i+1)$ contacts between Gln^{16} and Arg^{17} as well as Arg^{17} and Phe^{18}, indicating that this portion of the molecule is structured. These NOE's are compatible with a 3_{10} helix, but were weak and lacked the $d_{\alpha\beta}(i,i+3)$ NOE's to confirm the presence of a helix. The spectra observed for the C-terminus of NPAF may be due to a minor population of structures in solution that are nascent helices (30) or 3_{10} helices toward the end of the molecule. Also, the $^{3}J_{NH\text{-}H\alpha}$ coupling constants in this region of the molecule are conformationally ambiguous.

The family of lowest energy structures indicates the hydrogen bonding scheme for NPAF in TFE is abnormal. Two hydrogen bonds were detected in the helical region of the molecule. In 50% of the calculated structures, a usual alpha helical hydrogen bond was observed between NH of Ser^{11} and the CO Ser^{7}. In 85% of the structures, a hydrogen bond between Leu^{12} and Phe^{9} was observed.

SDS

SDS has been used as a membrane mimetic for membrane proteins, lipid proteins, and G-protein coupled receptor peptides, and has favourable properties for NMR structural studies as a membrane mimic. The micelle is relatively small (~35Å diameter) and spherical, resulting in reasonable correlation times and manageable line widths in an NMR spectrum (for reviews see refs. 31 and 32).

The concentration of SDS used for this study was 243 mM, well above the critical micelle concentration of 8mM (32). SDS is known to form micelles consisting of 60 SDS molecules (31,32); thus the micelle concentration was 4.1mM in this solution. At this micelle concentration and a peptide concentration of 3.0mM, we are assured that there is only one peptide molecule bound per micelle.

All spectra for the SDS solution were taken at 35°C where spectral overlap was least problematic. The elevated temperature allowed resolution of most backbone amide protons. It also resulted in narrower line widths more amenable for structural work and facilitated the measurement of $^3J_{HN-H\alpha}$ coupling constants from the amide region of the 1D spectrum.

The 1D spectrum of NPAF in SDS suggested that the molecule was structured, as there was a large chemical shift dispersion of the amide resonances (8.64-7.35) (vide infra). Also, several 1D spectra were taken at various temperatures, 25°C to 40°C, to deduce the $^3J_{HN-H\alpha}$ coupling constants for all amino acids. Line broadening at lower temperatures, due to the reduced correlation time of the micelle, and spectral overlap, which persisted at all temperatures, hampered

our efforts. However, for the amino acids where the $^3J_{HN-H\alpha}$ coupling constants could be measured, these were generally less than 6Hz (Table 4-2). Coupling constants less than the line width of the peak of interest are indicated as small. These data suggested that NPAF was adopting a stable helical conformation in the SDS solution, from residues Ser^7 to Trp^{10}. Furthermore, the chemical shift index for NPAF in SDS supported the alpha helical conformation (Fig. 4-5) within the central portion of the molecule, Ser^7 to Ser^{11}. These amino acids displayed consecutive up field H_α resonance signals. At 35°C, most of the amide resonances were well resolved except for pairs Leu^5 and Arg^{17}, and Ser^{11} and Ala^{13}. Additionally, there were a few Hα protons where spectral overlap was also problematic (Table 4-2 and Fig 4-6), so a higher magnetic field (800MHZ) was sought to remove the redundancy. Unfortunately, for the previously mentioned resonances, spectral overlap persisted at the higher magnetic field (Table 4-2 and Fig 4-6).

As observed in the TFE solution, NPAF in SDS displayed medium range NOE's indicative of a helical conformation (Figs. 4-6 and 4-7). A series of strong sequential $d_{NN}(i,i+1)$, medium range $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+3)$, $d_{\alpha N}(i,i+4)$, and $d_{\alpha\beta}(i,i+3)$ NOE's were present in the NOESY spectrum, also indicating that NPAF is adopting a helical conformation from Ser^7 to Ala^{14} (Fig. 4-7). The $d_{\alpha N}(i,i+4)$ strongly suggests that NPAF is adopting a stable alpha helical structure within the latter half of the helical region, Trp^{10} to Ala^{14}. It must be noted that due to spectral overlap, not all NOE's that were diagnostic of an alpha helix could be observed (Table 4-2 and Fig. 4-6). In addition, the $d_{\alpha\beta}(i,i+3)$ NOE's that were

observed were weak. As well, there were a few $d_{\alpha N}(i,i+2)$ NOE's that were observed within this region of the molecule. These data indicate that there may also be a small population of nascent helix and 3_{10} helix present in solution.

The helical region is comprised of the hydrophobic residues which form a hydrophobic cluster, as seen by several side chain interactions between hydrophobic amino acids in i,i+3 and i,i+4 positions. These were observed in the aromatic region of the spectra between the side chains of Phe^{9} and Leu^{12}, Phe^{9} and Ala^{13}, Trp^{10} and Ala^{13}, and finally Trp^{10} and Ala^{14} (data not shown).

The C-terminal tetrapeptide displays similar NOE information in SDS, as was observed in the TFE solution. In SDS, there were two consecutive $d_{\alpha N}(i,i+2)$ NOE's between Pro^{15} and Arg^{17}, and Gln^{16} and Phe^{18}. Also, a $d_{\alpha N}(i,i+3)$ was present in the spectra between Pro^{15} and Phe^{18}. A feature unique to the SDS solution was the hydrophobic contacts between the side chains of Phe^{18} and Ala^{14}, as well as an NOE between the β-protons of Ala^{14} and the NH of Phe^{18}. The $^{3}J_{HN\text{-}H\alpha}$ of Gln^{16} is less than 6Hz, indicating that the C-terminal tetrapeptide is conformationally stable.

To model NPAF in SDS, a total of 122 NOE based distances (32 intraresidue, 60 sequential, and 30 medium range restraints) and 7 dihedral angle restraints ($^{3}J_{NH\text{-}H\alpha}$<6Hz) were used. Two hundred structures were calculated in the modeling protocol. The twenty lowest energy conformers that satisfied all of the distance and dihedral angle restraints were chosen to represent the structure of NPAF when bound to SDS. The best fit superposition of the twenty NMR structures is shown in Figure 4-8A. Figure 4-8B indicates the pairwise RMSD for

each amino acid for the best fit superposition of Ser^7 to Ala^{14}. The RMSD for the family of conformers within the helical region of the molecule, Ser^7 to Ala^{14}, is 0.65 +/- 0.20Å for the backbone heavy atoms and 1.36 +/- 0.27Å for all heavy atoms within this region. The structural statistics for the modeling are given in Table 4-4.

The tertiary structure of NPAF when bound to SDS micelles can be described as helical, from Ser^7 to Ala^{14}. Ala^{14} defines the boundary of the helix and actually acts as the helix capping residue. The NMR evidence clearly indicates that NPAF is adopting a helical conformation, which is in agreement with the results of the molecular modeling.

The precise definition of the helix, whether 3_{10} or α-helix, is unclear from the NMR spectra; NPAF displays evidence for both kinds of helices, which are conformationally similar. In fact, the family of calculated structures reflects this result, as there is diversity in helicity, with structures displaying 3_{10} or α-helices. This result has been observed for a segment of the HIV glycoprotein gp41 when bound to dodecylphosphocholine micelles (DPC) (33) and porcine motilin bound to SDS micelles (34).

The C-terminal tetrapeptide forms a type I β-turn centered on Gln^{16} and Arg^{17} (Fig. 4-8C). The average ϕ/ψ dihedral angles at these amino acids are -54.0° +/- 1.9° and -23.7° +/- 1.5° for Gln^{16}, and -76.8° +/- 15.6° and -13.0° +/- 17.1° for Arg^{17}, respectively. The accepted ϕ/ψ values for a type I β-turn are -60° and -30°, and -90° and 0° for the residues in the i+1 and i+2 positions of the β-turn, respectively (26). The NMR evidence is also compatible with a type I β-turn. The

expected NOE's for a type I β-turn (26) consist of a medium range $d_{\alpha N}(i,i+2)$ and $d_{NN}(i,i+2)$ NOE between residues i+1 and i+3 of the turn; a $d_{\alpha N}(i,i+3)$ between the first and last amino acids of the turn; an intense $d_{NN}(i,i+1)$ NOE between of the amide protons of the third and fourth amino acids; a $d_{NN}(i,i+1)$ NOE between of the amide protons of the second and third amino acids of the turn. Both a 4 Hz and 9 Hz $^3J_{HN\text{-}H\alpha}$ coupling constants are expected at the second and the third residues of the turn, respectively (26). Often a hydrogen bond is seen in β-turns between the amide proton of the fourth residue and the carbonyl of the first amino acid. The NOE's that were observed for NPAF (Fig. 4-7) are compatible with the above description, and the ~5Hz $^3J_{HN\text{-}H\alpha}$ coupling constant seen at Gln^{16} is also compatible with a type I β-turn. Unfortunately, a $^3J_{HN\text{-}H\alpha}$ could not be obtained for Arg^{17} due to spectral overlap, however the NOE evidence does suggest a β-turn. In addition, the hydrophobic contacts between Ala^{14} and Phe^{18} are across the turn and place Phe^{18} in the correct orientation for a β-turn. Finally, for 95% of the calculated structures, a hydrogen bond was present between Pro^{15} and Phe^{18}.

The type I β-turn is preceded by an inverse gamma turn centered on Pro^{15}. An inverse gamma turn is defined as a three residue turn, characterized by the ϕ/ψ angles of the central residue (i+1 amino acid of the turn). The accepted ϕ/ψ values for an inverse gamma turn are -75° and 69° (35,36). The observed value of the ϕ/ψ dihedral angles for Pro^{15} are -61.6° +/- 7.0° and 118.6° +/- 3.3°, respectively, indicating that the turn is somewhat distorted.

Alternatively, the C-terminal structure has some of the features of a proline C-capping motif (37, 38). The nomenclature for helix capping is as follows: C3-

C2-C1-Ccap-C1'-C2'-C3'-C4', where the numbered values indicate residues within the helix, Ccap is the helix capping residue, and the primed values indicate residues outside of the helix. The proline capping motif is described sequentially by a proline at the C' position, and a hydrophobic residue at the C4' position. Positions C2' and C3' are occupied by polar residues. Sequentially, this motif is satisfied by NPAF and the corresponding residues are Ser^{11}-Leu^{12}-Ala^{13}-Ala^{14}-Pro^{15}-Gln^{16}-Arg^{17}-Phe^{18}. Structurally, a bifurcated hydrogen bond between the amide protons of C3' and C4' and the carbonyl oxygen at the C-capping residue, and a hydrophobic interaction between C4' and C2, are expected (37). Idealized dihedral angles have also been indicated for the last five amino acids in the capping motif (37), and distortions at the Ccap and C1' positions have been noted (38). The average ϕ/ψ dihedral angles for Ala^{14} are -70.9° +/- 11.8° and 150.9° +/- 8.6°, respectively; the average ϕ/ψ dihedral angles for Pro^{15} were listed above. However, the results from our modeling do not fit within the bounds listed by Aurora and Rose (37), nor was a hydrophobic contact observed between Phe^{18} and Leu^{12} (C4' and C2) in our spectra. In addition, our modeling does not indicate a bifurcated hydrogen bond involving Ala^{14}, the C-capping residue.

Spin Label Studies

5-DS was added to the SDS sample to examine the placement of NPAF in SDS. 5-DS contains an unpaired electron that induces paramagnetic relaxation resulting in line broadening of resonance signals in proximity to the spin label. Addition of 5-DS to the sample resulted in line broadening of all resonances within the spectra of NPAF, as judged by the 1D spectra (Fig. 4-9). A comparison

of the cross peak intensities in the TOCSY spectra with and without the spin label (Fig. 4-10) also supports this conclusion. The line broadening that is observed across the spectrum indicates that NPAF is associating with the micelle. The residues that were affected the most by the spin label are within the hydrophobic cluster of the helix (Phe^9 to Ala^{14}), as well as Leu^5. 5-DS probes the surface of the micelle (31,39,40), so it is likely that the NPAF is sitting on the surface of the micelle, bound through hydrophobic interactions between the hydrophobic core of the micelle and the hydrophobic cluster of NPAF. Leu^5 was also affected by the spin label and is outside of the helical region of the molecule. Thus Leu^5 acts as an additional anchor to bind NPAF to the micelle surface. Long chain aliphatic amino acids have been used to increase the binding affinity of peptides to micelles through hydrophobic interactions with the micelle (41). The side chain protons of Trp^{10}, an amino acid known to anchor peptides to membrane surfaces by having a preference for water membrane surface interface (42,43), were affected the most by the spin label (Fig. 4-9), indicating that Trp^{10} is in closest proximity to the spin label.

Summary and Conclusions

The membrane associating properties of NPAF are consistent with the Membrane Compartments Theory proposed by Schwyzer (9,10,11). According to this theory, the target cell membrane induces preferred conformations and orientations of peptides by guiding important residues into the hydrophobic, fixed charge, or aqueous environments of the cell membrane. Our results are consistent with this model, since in SDS, the micelle induces a unique feature: the C-

terminal tetrapeptide is conformationally stable. In both solvent systems the central part of the molecule consists of an α-helix, specifically Asn^6 to Ala^{13} in TFE/H_2O solution, and Ser^7 to Ala^{14} in SDS solution. The helical region contains a hydrophobic cluster, which serves to bind NPAF to the micelle, with an additional anchor Leu^5, just outside the helical region, as seen in our spin label studies.

A more detailed comparison of the data in both solvent systems indicates that in going from TFE/H_2O solutions to SDS micellar solutions, the C-terminal tetrapeptide is much more structured in the latter, consisting of an inverse γ-turn (residues Ala^{14} to Gln^{16}) and a β-turn involving residues Pro^{15} and Phe^{18}. The change in conformation suggests that the last four amino acids are important in receptor recognition. In fact, it has been shown for other small G-protein coupled receptor peptides that a change in conformation upon binding to a membrane mimetic is likely to be important for receptor recognition. For example, the small opiate peptides such as methionine and leucine enkephalins form β-turns when bound to membrane mimetics (44, 45, 46) as do bradykinin agonists and antagonists (47). Bradykinin is a potent vasodilator and has analgesic effects (47). Interestingly, it has been shown recently by substitution studies for the related peptide NPFF that the C-terminal tripeptide, QRF-NH_2, is important for receptor affinity and activity on the human NPFF2 or HLWAR77 receptor (48). NPFF and NPAF share the same receptors and C-terminal tetrapetide sequence, PQRF-NH_2.

An additional feature of the Membrane Compartments Theory describes the process that occurs after membrane association and pre-folding. The peptide

bound to the membrane surface is free to diffuse two-dimensionally along the membrane surface to the receptor, where receptor recognition and binding occurs. On the basis of the micelle associating properties of NPAF seen in this study, it is likely that NPAF undergoes this two-step model. This structural data may lead to the knowledge that will facilitate the identification of receptor agonists, antagonists dealing with pain modulation and opioid tolerance, as well new drug discoveries for the treatment of pain and drug addiction.

Table 4-1: The chemical shifts (ppm) and coupling constants (Hz) for NPAF in 50%/50% TFE/H_2O at 20°C.

Amino Acid	NH	α(α')	β(β')	γ(γ')	δ(δ')	Other	$^3J_{HN-H\alpha}$
Ala1	-	4.23	1.57				-
Gly2	8.75	4.12(3.76)					-
Glu3	9.15	4.15	2.08(2.00)	2.31			4.3
Gly4	8.48	3.99(3.96)					5.6
Leu5	7.90	4.33	1.73	1.62	0.91(0.84)		-
Ans6	8.19	4.72	2.92			Hδ-7.52(6.76)	-
Ser7	8.20	4.23	3.98(3.91)				4.5
Gln8	8.41	4.20	2.02	2.30		Hε-7.17(6.64)	5.1
Phe9	8.01	4.33	3.06			Hδ1,2-6.89	4.8
Trp10	7.72	4.47	3.41(3.33)			Hδ1-7.31 Hε1-9.86 Hε3-7.58 Hζ2-7.50 Hζ3-7.08 Hη2-7.16	3.8
Ser11	7.86	4.28	3.98(3.80)				5.5
Leu12	7.59	4.31	1.73	1.55	0.91(0.88)		-
Ala13	7.82	4.27	1.27				-
Ala14	7.81	4.45	1.40				-
Pro15	-	4.39	2.36(1.93)	2.06	3.75(3.66)		-
Gln16	8.26	4.25	2.14(2.05)	2.42		Hε-7.37(6.61)	6.1
Arg17	7.96	4.20	1.65	1.40	3.07	εNH-7.08	6.5
Phe18	7.89	4.66	3.30(3.00)			Hδ1,2-7.32	-
NH2	7.24(6.92)						-

Table 4-2: The chemical shifts (ppm) and coupling constants (Hz) for NPAF in 243mM SDS at 35°C

Amino Acid	NH	α(α')	β(β')	γ(γ')	δ(δ')	Other	$^3J_{HN-H\alpha}$
Ala1	-	4.24	1.56				-
Gly2	8.58	4.10(3.90)					-
Glu3	8.64	4.27	2.08(1.99)	2.36(2.30)			5.4
Gly4	8.37	3.99					5.7
Leu5	7.87	4.31	1.72	1.56	0.88(0.81)		-
Ans6	8.22	4.69	2.88(2.84)			Hδ-7.49(6.80)	6.1
Ser7	8.06	4.28	3.95(3.84)				3.9
Gln8	8.28	4.11	1.93	2.24		Hε-7.27(6.71)	5.6
Phe9	7.93	4.27	3.01(2.89)			Hδ1,2-6.82	Small
Trp10	7.34	4.44	3.26(3.14)			Hδ1-7.25 Hε1-9.91 Hε3-7.40 Hζ2-7.44 Hζ3-7.02 Hη2-7.13	Small
Ser11	7.75	4.34	3.89(3.84)				-
Leu12	7.51	4.26	1.70	1.52	0.89(0.87)		-
Ala13	7.75	4.27	1.26				-
Ala14	7.83	4.48	1.35				4.7
Pro15	-	4.46	2.35(1.94)	2.01	3.69(3.63)		-
Gln16	8.26	4.22	2.15(2.08)	2.44(2.39)		Hε-7.45(6.77)	5.3
Arg17	7.87	4.12	1.60(1.45)	1.26(1.22)	3.01	εNH-6.98	-
Phe18	7.67	4.66	3.33(2.89)			Hδ1,2-7.31	8.3
NH2	7.28(6.98)						-

Table 4-3: The structural statistics for NPAF in TFE/H_2O.

Energy

	Average	Standard Deviation
E_{Tot}(kcal/mol)	14.57	0.26
E_{VDW}(kcal/mol)	1.57	0.04
E_{NOE}(kcal/mol)	0.83	0.14

Violations

	RMSD	Standard Deviation
Bond	0.0010	0.00008
Angle	0.3951	0.0017
Improper	0.0821	0.0027
NOE	0.0101	0.0009
Dihedral	0.0023	0.0057

Table 4-4: The structural statistics for NPAF in SDS.

Energy

	Average	Standard Deviation
E_{Tot}(kcal/mol)	15.50	0.26
E_{VDW}(kcal/mol)	2.52	0.18
E_{NOE}(kcal/mol)	0.63	0.21

Violations

	RMSD	Standard Deviation
Bond	0.0009	0.00006
Angle	0.3979	0.0023
Improper	0.0873	0.0037
NOE	0.0082	0.0013
Dihedral	0.1878	0.0389

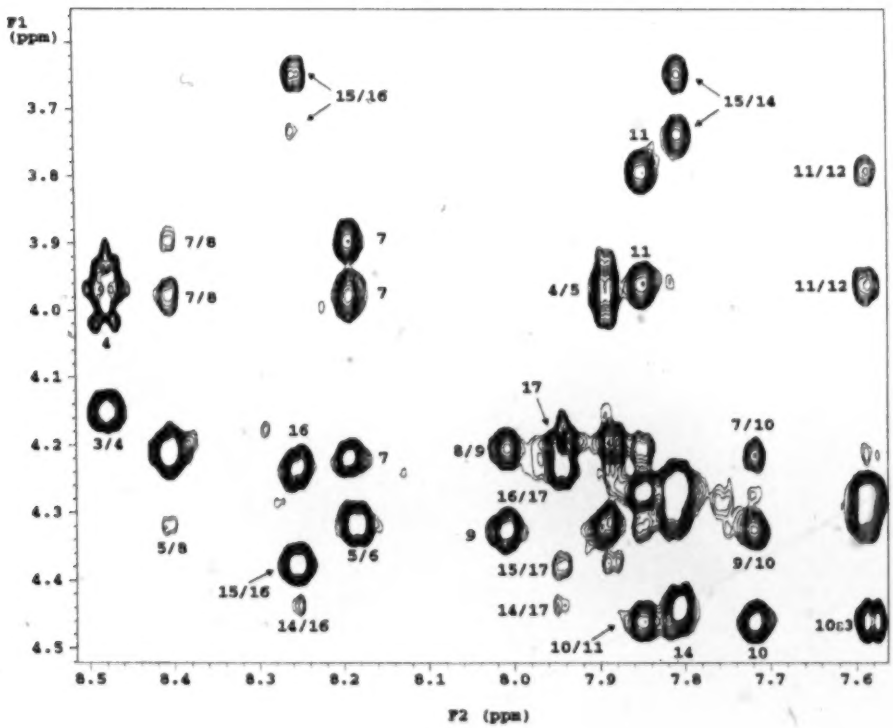


Figure 4-1. Part of the NH-Hα region of the 600MHz NOESY with a 150ms mixing time for NPAF in 50%/50% TFE/H_2O. Assignments are indicated.

Figure 4-2: The summary of important NOE data for NPAF in 50%/50% TFE/H_2O at 20°C. The bars indicate the observed NOE between the amino acids involved. The thickness of the bar for $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ indicate the strength of the NOE. The arrows indicate an amino acid with a $^3J_{HN\text{-}H\alpha}$ less than 6Hz.

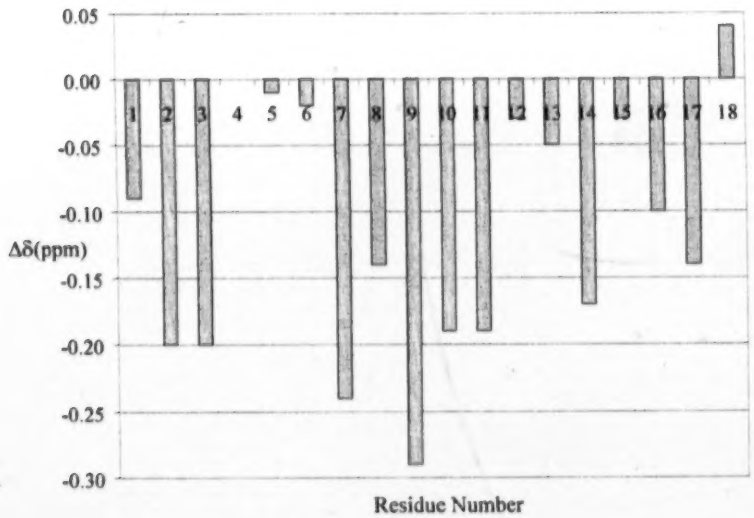
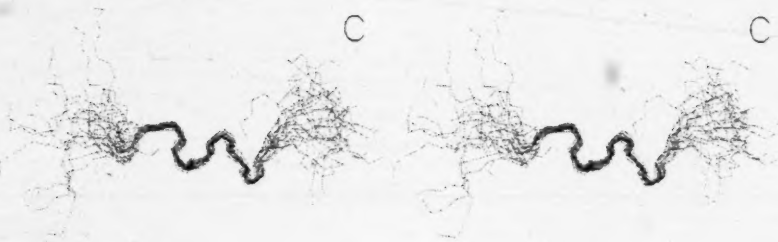


Figure 4-3. The H_α chemical shift index of NPAF in TFE/H_2O.

A)



B)

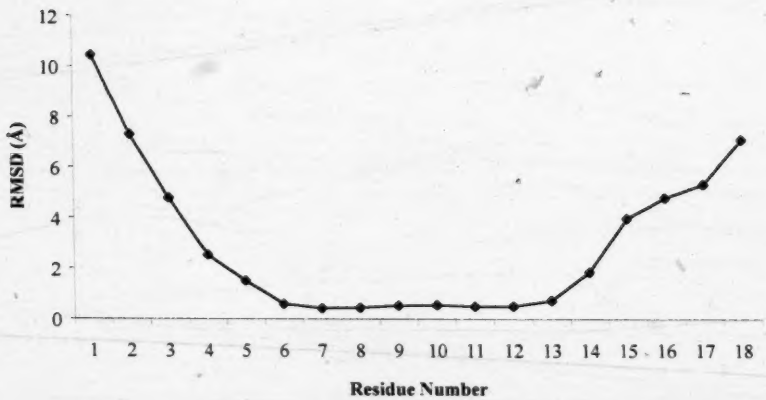


Figure 4-4. The stereo view of the twenty lowest energy structures of NPAF in TFE/H_2O showing the best fit superposition of residues Asn^6 to Ala^{13}. The RMSD for the family of conformers from Asn^6 to Ala^{13} for the backbone heavy atoms is 0.56 +/- 0.17Å and for all heavy atoms is 1.50 +/- 0.28Å. The C-terminus of the molecule is indicated. B) The global pairwise per residue RMSD for all backbone heavy atoms for NPAF in TFE/H_2O, as calculated by Molmol.

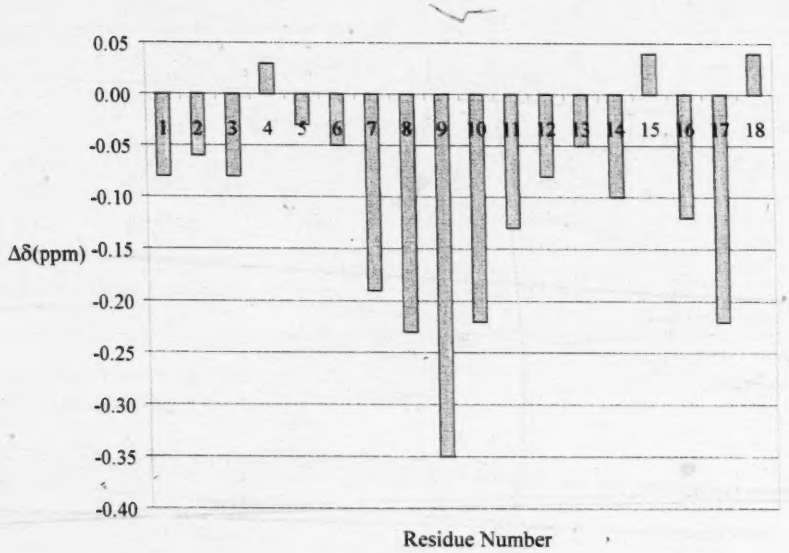


Figure 4-5. The chemical shift index for NPAF in SDS.

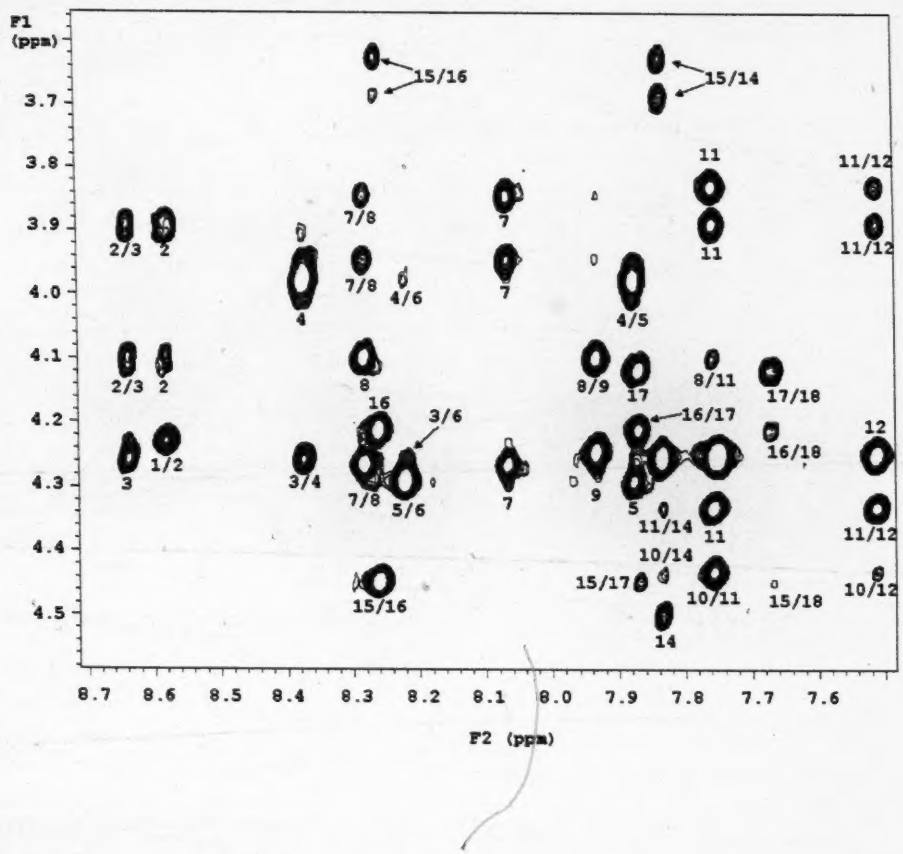
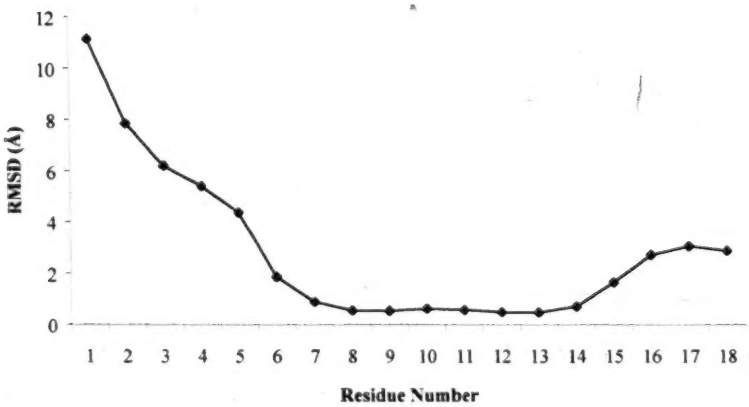


Figure 4-6. Part of the NH-Hα region of the 150ms mixing time 800MHz NOESY spectrum of NPAF in 243mM SDS taken at 35°C. Assignments are indicated.

Figure 4-7. The summary of important NOE data for NPAF in 243mM SDS at 35°C. The bars indicate the observed NOE between the amino acids involved. The thickness of the bar for $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ indicate the strength of the NOE. The arrows indicate an amino acid with a $^3J_{HN\text{-}H\alpha}$ less than 6Hz.

A)

B)



C)

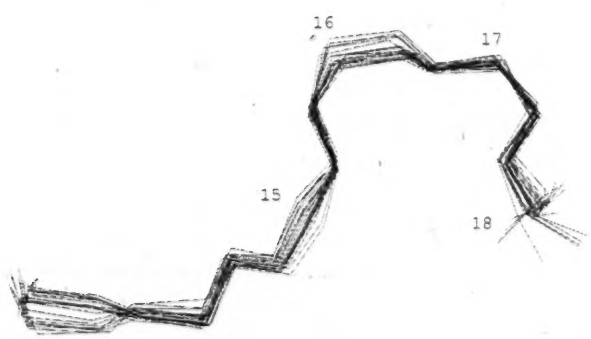


Figure 4-8. A) The stereo view of the twenty lowest energy structures of NPAF in 243mM SDS showing the best fit superposition of residues Ser^{7} to Ala^{14}. The RMSD for the family of conformers from Ser^{7} to Ala^{14} for the backbone heavy

atoms is 0.65 +/- 0.20Å and for all heavy atoms is 1.36 +/- 0.27Å. The C-terminus of the molecule is indicated. B) The global pair wise per residue RMSD for all backbone heavy atoms for NPAF bound to SDS micelles, as calculated by Molmol. C) The best fit superposition of the C-terminal hexapeptide, residues Ala^{13} to Phe^{18} emphasizing the type I β-turn. The pairwise RMSD for all backbone heavy atoms for residues Ala^{13} to Phe^{18} is 0.50 +/- 0.23Å.

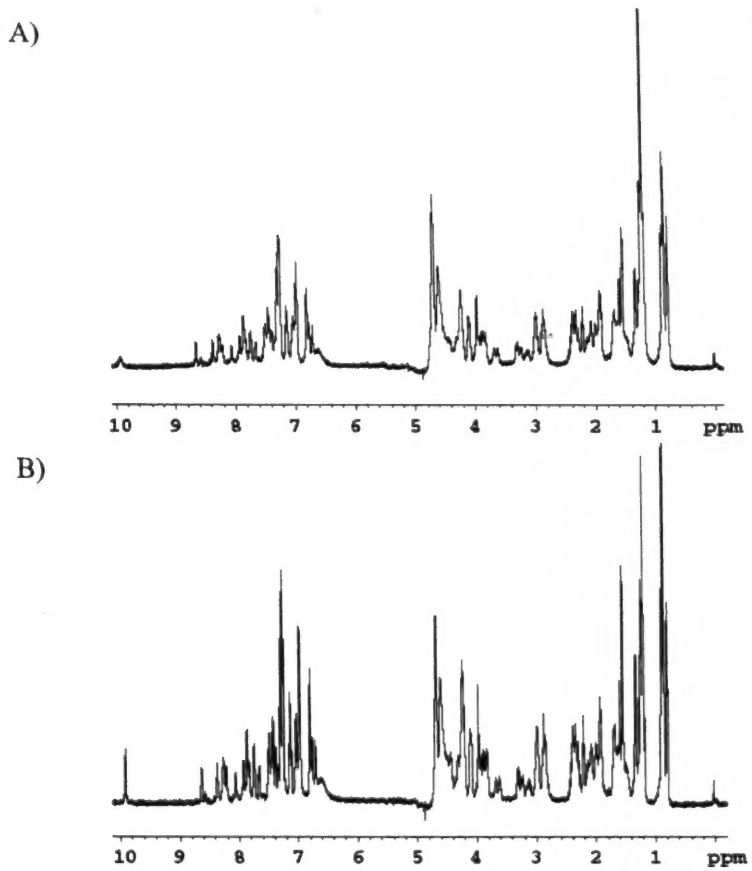


Figure 4-9. One-dimensional spectra of NPAF in the SDS solution: A) in the presence and B) in the absence of the spin label 5-doxyl-stearic acid. The spectra were recorded in a similar manner on the Varian Unity 500MHz spectrometer at 35°C and the water peak was attenuated by presaturation. No weighting functions were added to the spectra and the residual water peak was further attenuated by the addition of a digital filter. Spectra are shown using the same vertical scale.

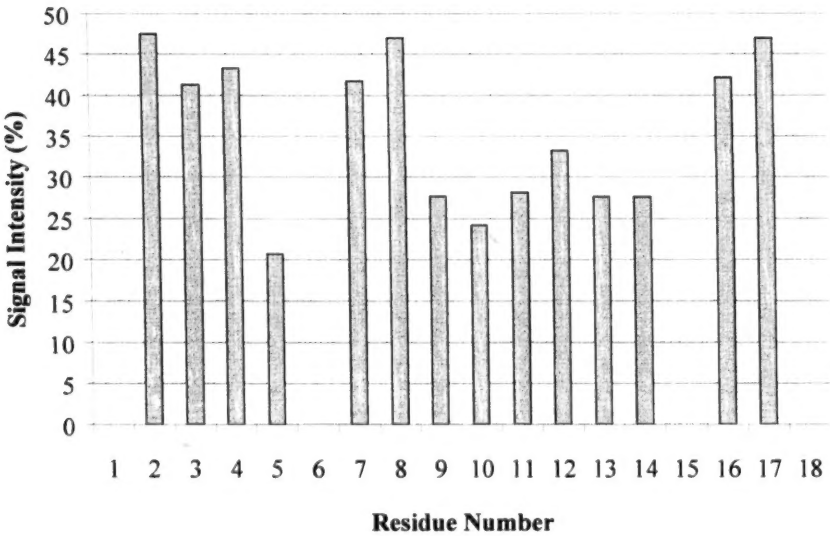


Figure 4-10. The percentage of signal intensity remaining after the addition of the 5-DS for HN-Hα TOCSY cross peaks. The residual water peak obscured the Hα resonance signals of Asn^{6} and Phe^{18}, so these data were not included, and the average value is given for Gly^{2} Hα protons.

References

1. H.-Y.T. Yang, W. Fratta, E.A. Majane, and E. Costa, *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 7757-7761 (1985).

2. S.J. Perry, E.Y.-K. Huang, D. Cronk, J. Bagust R. Sharma, R.J. Walker, S. Wilson, and J.F. Burke, *FEBS Lett.*, **409**, 426-430 (1997).

3. P. Panula, A.A. Aarnisalo, and K. Wasowicz, *Prog. Neurobiol.*, **48**, 461-487 (1996).

4. M. Roumy, and J.M. Zajac, *Eur. J. Pharmacol.*, **345**, 1-11 (1998).

5. E.S. Vilim, and E. Ziff, *Soc. Neurosci. Abstr.*, **21**, 760 (1995).

6. N.A. Elshourbagy, R.S. Ames, L.R. Fitzgerald, J.J. Foley, J.K. Chambers, P.G. Szekeres, N.A. Evans, D.B. Schmidt, P.T. Buckley, G.M. Dytko, P.R. Murdock, G. Milligan, D.A. Groarke, K.B. Tan, U. Shabon, P. Nuthulaganti, D.Y. Wang, S. Wilson, D.J. Bergsma, and H.M. Sarau, *J. Biol. Chem.*, **275**, 25965-25971, (2000).

7. J.A. Bonini, K.A. Jones, N. Adham, C. Forray, R. Artymyshyn, M.M. Durkin, K.E. Smith, J.A. Tamm, L.W. Boteju, P.P. Lakhlani, R. Raddatz, W.-J. Yao, K.L. Ogozalek, N. Boyle, E.V. Kouranova, Y. Quan, P.J. Vaysse, J.M. Wetzel, T.A. Branchek, C. Gerald, and B. Borowsky, *J. Biol. Chem.*, **275**, 39324-39331 (2000).

7. R. Bader, A. Bettio, A.G. Beck-Sickinger, and O. Zerbe, *J. Mol. Biol.*, **305**, 307-329 (2001).

8. R. Bader, A. Bettio, A.G. Beck-Sickinger, and O. Zerbe, *J. Mol. Biol.*, **305**, 307-329 (2001).

9. R. Schwyzer, *Biochemisty*, **25**, 6335-6342 (1986).

10. R. Schwyzer, *Biopolymers*, **31**, 785-792 (1991).

11. R. Schwyzer, *Biopolymers*, **37**, 5-16 (1995).

12. J.D.J O'Neil, and B.D. Sykes, *Biochemistry*, **28**, 699-707 (1989).

13. A. Bax, and D.G. Davis, *J. Magn. Reson.*, **65**, 355-360 (1985).

14. M. Levitt, R. Freeman, and T. Frankeil, *J. Magn. Reson.*, **47**, 328 (1982).

15. R.E. Hurd, *J. Magn. Reson.*, **87**, 422-428 (1999).

16. M. von Kienlin,.C.T.W. Moonen, A. van der Toorn, and P.C.M. van Zijl, *J. Magn. Reson.*, **93**, 423-429 (1991).

17. S. Macura, and R.R. Ernst, *Mol. Phys.*, **41**, 95-117 (1980).

18. D.J. States, R.A. Haberkorn, and D.J. Ruben, *J. Magn. Reson.*, **48**, 286-292 (1982).

19. B.A. Johnson, and R.A. Blevins, *J. Biomol. NMR.*, **4**, 603-614, (1994).

20. M. Nilges, A.M. Gronenborn, A.T. Brünger, and G.M. Clore, *Protein Eng.*, **2**, 27-38 (1988).

21. A.T. Brünger; *X-PLOR. Version 3.1: A system for X-ray crystallography and NMR*; Yale University Press: New Haven, USA, 1993.

22. M. Nilges, *Proteins: Structure, Function and Genetics*, **17**, 297-309 (1993).

23. C.M. Fletcher, D.N.M. Jones, R. Diamond, and D. Neuhaus, *J. Biomol. NMR.*, **8**, 292-310 (1996).

24. A.T. Brünger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, and G.L. Warren, *Acta. Cryst.*, **D54**, 905-921 (1998).

25. R. Koradi, M. Billeter, and K. Wüthrich, *J. Mol. Graph.*, **14**, 51-55 (1996).

26. K. Wüthrich, *NMR of Proteins and Nucleic Acids*; John-Wiley and Sons: New York, 1986.

27. D.S. Wishart, C.G. Bigam, A. Holm, R.S. Hodges, and B.D. Sykes, *J. Biomol. NMR.*, **5**, 67-81 (1995).

28. D.S. Wishart, B.D. Sykes, and F.M. Richards, *J. Mol. Biol.*, **222**, 311-333 (1991).

29. D.S. Wishart, B.D. Sykes, and F.M. Richards, *Biochemistry*, **31**, 1647-1651 (1992).

30. H.J. Dyson, M. Rance, R.A. Houghten, P.E. Wright, and R.A. Lerner, *J. Mol. Biol.*, **201**, 201-217 (1988).

31. P. Damberg, J. Jarvet, and A. Gräslund, *Methods Enzymol.*, **339**, 271-285 (2001).

32. G.D. Henry and B.D. Sykes, *Methods Enzymol.*, **239**, 515-535 (1994).

33. D.J. Schibli, R.C. Montelaro, and H.J. Vogel, *Biochemistry*, **40**, 9570-9578 (2001).

34. J. Jarvet, J. Zdunek, P. Damberg, and A. Gräslund, *Biochemistry*, **36**, 8153-8163 (1997).

35. E.J. Milnerwhite, E.J.; B.M. Ross, B.M.;R. Ismail, K. Belhadjmostefa, and R. Poet R. *J. Mol. Biol.*,**204**, 777-782 (1988).

36. G.D. Rose, L. Gierasch, L., and J.A. Smith, *Adv. Protein. Chem.*, **37**, 1-109, (1985).

37. R. Aurora, and G.D. Rose, *Protein Sci.*, **7**, 21-38 (1998).

38. J. Prieto and L. Serrano, *J. Mol. Biol.*, **274**, 276-288 (1997).

39. L.R. Brown, C. Bösch, and K. Wüthrich, *Biochim. Biophys. Acta.*, **642**, 296-312 (1981).

40. C.H.M. Papavoine, R.N.H Konings, C.W. Hilbers, and F.J.M. van de Ven, *Biochemistry*, **33**, 12990-12997 (1994).

41. A.M. Dixon, R.M. Venable, R. Pastor, and T.E. Bull, *Micelle-Bound Conformations of a Peptide from Hemagglutinin A: Stabilization of a C-terminal Beta Turn*; April 2002, 43rd ENC, Asilomar CA, Poster no. 209.

42. W.-M. Yau, W.C. Wimley, K. Gawrisch, and S.H. White, *Biochemistry*, **37**, 14713-14718 (1998).

43. W.C. Wimley, and S.H. White, *Nat. Struct. Biol.*, **3**, 842-848 (1996).

44. B.A. Behnam, and C.M. Deber, *J. Biol. Chem.*, **259**, 14935-14940 (1984).

45. A. Milon, T. Miyazawa, and T. Higashijima, *Biochemistry*, **29**, 65-75 (1990).

46. I. Marcotte, S. Gagne, F. Separovic, and M. Auger, *The Conformation of Met-Enkephalin in Interaction with Zwitterionic and Anionic Bicelles: a Multidimensional ^{1}H NMR Study*; August 2002, 20th International Conference on Magnetic Resonance in Biological Systems, Toronto Ont., Poster no. 262.

47. G. Kotovych, J.R. Cann, J.M. Stewart, H. Yamamoto, *Biochem. Cell. Biol.*, **76**, 257-266 (1998).

48. H. Mazarguil, C. Gouarderes, J.A.M. Tafani, D. Marcus, M. Kotani, C. Mollereau, M. Roumy, and J.M. Zajac, *Peptides*, **22**, 1471-1478 (2001).

Chapter 5.

Concluding Chapter

Conclusions

Throughout this thesis the secondary and tertiary structure of small peptides were examined by ^{1}H 2D NMR. The structural features of the molecules studied in this thesis were examined in an effort to gain structural/functional relationships for the peptides. It must be noted that the choice of solvent was of importance for these studies, an idea that was emphasized almost 30 years ago (1) and has been re-emphasized recently in a study of the disordered protein FlgM (negative regulator of flagellim synthesis) (2), studied at high co-solute concentrations. The peptides that were studied in this thesis are G-protein coupled receptor peptides that are postulated to associate with the cell membrane, where the membrane guides areas of the peptide into the hydrophilic and hydrophobic compartments of the membrane, thus resulting in pre-folding of the peptide. The peptide then undergoes two-dimensional diffusion along the membrane surface to the receptor where peptide recognition and binding occur (3). As a result, the solvents chosen in this thesis were intended to mimic the cell membrane to induce peptide structure (4,5).

In Chapter 2, the secondary structure of B-10324 was determined. The secondary structure of B-10324 consisted of a distorted type II β-turn from residues Pro^2 to Cpg^5 with a stabilizing hydrogen bond between the first and last amino acids of the turn, Pro^2 and Cpg^5. The structure of B-10324 was found to be similar to other BK B_1 antagonists, B-9858 (6) and B-10148 (7), which also had a 2-5 type II β-turn. In analogy with B-9958, a BK B_1 antagonist having the same amino acid sequence as B-10324 but lacking the N-capping F5C group, the

structural properties of B-9858, B-10148, and B-10324 were examined in the context of B_1 activity measurements (8). The series of peptides were structurally similar since all displayed an N-terminal 2-5 type II β-turn but differed in activity depending on the number of Igl residues. B-9858 and B-10148 were non-competitive, nonequilibrium antagonists having two and one Igl residues, respectively, with B-9858 having greater activity. B-10324 was a surmountable antagonist with no Igl residues. The difference in activity between the molecules was attributed to the presence and number of Igl residues.

Chapter 3 of this thesis presented research on an invertebrate peptide from the parasitic tape worm *Moniezia expansa*, namely NPF. The secondary structure of NPF was found to be very similar to the related vertebrate peptide porcine NPY. The secondary structure of NPF was found to consist of an amphipathic α-helix from Lys^{14} to Ile^{31}, which is in a similar region as found in NPY. The amphipathic α-helix found in NPY, when bound to DPC micelles, is from Ala^{14} to Tyr^{36} (9). Based on the similarities between these two peptides, we speculate that the mode of action for NPF will be consistent with the Membrane Compartments Theory (3).

Finally, in Chapter 4 the secondary structure of NPAF was determined in two different solvent conditions, in a TFE/H_2O solution, and when bound to SDS micelles. The secondary structure of NPAF in both solvent conditions was very similar as it was predominantly alpha helical within the central portion of the molecule, from approximately Ser^7 to Ala^{14}. With a change in solvent from TFE/H_2O to SDS, a type I β-turn was induced at the C-terminus of NPAF. In

combination with substitution studies (10), which have shown that the C-terminal tripeptide is important for receptor affinity and activity, these data emphasize the importance of the C-terminus for receptor affinity and activity. The membrane associating properties of NPAF indicate that its mode of action is consistent with the Membrane Compartments Theory.

Future Work

The conformations of B-10324, NPF, and NPAF by 2D ^{1}H NMR presented in this thesis are not the final word on conformational studies for these molecules but a beginning. Indeed, NMR can be employed further to investigate the conformations of the molecules studied in this thesis. Presented below is a brief outline of future work that can be carried out on the molecules studied in this thesis.

B-10324

Many BK related peptides have been synthesized and structural studies have been performed (5). The structural features of B-10324 are consistent with these studies. Efforts are still ongoing to apply these molecules as potential drug candidates based on the results of structural studies (11).

NPF

NPF's continue to be a source of interest (12) even though a specific function has not been applied to any of these peptides (12). The amphipathic helix of NPF suggests that it may undergo membrane association. It would be instructive to exam NPF under such conditions, when bound to SDS or DPC micelles, to gain insight into structural changes that may be of biological

relevance upon membrane binding. In addition, the placement of a peptide in or on a micelle can also be of biological relevance. To investigate this aspect of NPF when bound to a micelle, spin labels can be added to the micelle solution. Ideally, two spin labels can be used, namely 5-doxyl stearic acid and 16-doxyl stearic acid, which can be added to two different solutions. The nitroxyl group of the fatty acid is simply at carbon 5 or 16 of the fatty acid, respectively, and each fatty acid will probe different depths of the micelle (4) and indicate the placement of the peptide on the surface or inside the micelle.

NPAF

NPAF has been indicated as an important peptide in energy regulation, specifically adipocyte metabolism (13), emphasizing the importance of this peptide. As an extension to the structural study presented in chapter 4 of this thesis, it would be of interest to study the dynamic properties of NPAF, in particular backbone dynamics through ^{15}N relaxation measurements (14,15 and references therein). The motional properties of a molecule can be used to gain insight into structural features of the molecule. To meet this end, relaxation measurements in conjunction with the model free formalism of Lipari-Szabo (16,17) for the spectral density function can be performed. Pulse sequences are available (14,15,18) and software (19) is also available to aid in the analysis of the data. The peptide should be ^{15}N enriched at all amide nitrogens to fully explore the dynamic properties of the molecule at all positions and aid in the sensitivity. Alternatively, a similar approach can be performed using natural abundance ^{13}C

relaxation of the C_α or C' positions at elevated peptide concentrations to ensure high enough sensitivity for the measurements (18).

References

1. C.B. Anfinsen, *Science*, **181**, 223-230 (1973).

2. M.M. Dedmon, C.N. Patel, G.B. Young, and G.J. Pielak, *Proc. Natl. Acad. Sci. USA*, **99**, 12681-12684 (2002).

3. R. Schwyzer, *Biopolymers*, **37**, 5-16 (1995).

4. P. Damberg, J. Jarvet, and A. Gräslund, *Methods Enzymol.*, **339**, 271-285 (2001).

5. G. Kotovych, J.R. Cann, J.M. Stewart and H. Yamamoto, *Biochem. Cell. Biol.*, **76**, 257-266 (1998).

6. J. Sejbal, Y. Wang, J.R. Cann, J.M. Stewart, L. Gera and G. Kotovych, *Biopolymers*, **42**, 521-535 (1997).

7. M. Miskolzie, L. Gera, J.M. Stewart and G. Kotovych, *J. Biomolec. Struct. Dyn.*, **18**, 249-260 (2000).

8. J.-F. Larrivée, L. Gera, S. Houle, J. Bouthillier, D.R. Bachvarov, J.M. Stewart and F. Marceau, *Br. J. Pharmacol.*, **131**, 885-892 (2000).

9. R. Bader, A. Bettio, A.G. Beck-Sickinger, and O. Zerbe, *J. Mol. Biol.*, **305**, 307-329 (2001).

10. H. Mazarguil, C. Gouarderes, J.A.M. Tafani, D. Marcus, M. Kotani, C. Mollereau, M. Roumy, and J.M. Zajac, *Peptides*, **22**, 1471-1478 (2001).

11. Personal communication with Vellarkad Viswanadhan of Amgen.

12. D. Stanek, J. Pohl, J. Crim, and M. Brown, *Peptides*, **23**, 1367-1378 (2002).

13. I. Lefrère, P. De Coppet, J.-C. Camelin, S. Le Lay, N. Mercier, N. Elshourbagy, A. Bril, I. Berrebi-Bertrand, B. Fève, and S. Krief, *J. Biol. Chem*, **277**, 39169-19178 (2002).

14. J.H. Viles, D. Donne, G. Kroon, S.B. Prusiner, F.E. Cohen, H.J. Dyson, and P.E. Wright, *Biochemistry*, **40**, 2743-2753 (2001).

15. P.F. Flynn, R.J. Bieber Urbauer, H. Zhang, A.L. Lee, and A.J. Wand, *Biochemistry*, **40**, 6559-6569 (2001).

16. G. Lipari, and A. Szabo, *J. Am. Chem. Soc.*, **104**, 4546-4559 (1982).

17. G. Lipari, and A. Szabo, *J. Am. Chem. Soc.*, **104**, 4559-4570 (1982).

18. A.G. Palmer, M. Rance, and P.E. Wright, *J. Chem. Soc.*, **113**, 4371-4380 (1991).

19. Modelfree version 4.15, freely available from A.G. Palmer on the world wide web at:

http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/modelfree.html